GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:32:18 ; Search time 14.3333 Seconds
(without alignments)
994.100 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1522	100.0	276	4	US-09-094-964-1	Sequence 1, Appli
2	1522	100.0	348	3	US-08-652-265-2	Sequence 2, Appli
3	1522	100.0	348	3	US-08-834-497A-2	Sequence 2, Appli
4	1522	100.0	348	3	US-09-503-444A-2	Sequence 2, Appli
5	1522	100.0	348	4	US-09-277-457-2	Sequence 2, Appli
6	1522	100.0	348	4	US-09-679-729-2	Sequence 2, Appli
7	1513	99.4	276	4	US-09-094-964-2	Sequence 2, Appli
8	1513	99.4	348	3	US-08-652-265-6	Sequence 6, Appli
9	1513	99.4	348	3	US-08-834-497A-6	Sequence 6, Appli
10	1513	99.4	348	3	US-09-503-444A-6	Sequence 6, Appli
11	1511	99.3	348	3	US-08-652-265-4	Sequence 4, Appli
12	1511	99.3	348	3	US-08-834-497A-4	Sequence 4, Appli
13	1511	99.3	348	3	US-09-503-444A-4	Sequence 4, Appli
14	1502	98.7	276	4	US-09-094-964-3	Sequence 3, Appli
15	1502	98.7	348	3	US-08-652-265-8	Sequence 8, Appli
16	1502	98.7	348	3	US-08-834-497A-8	Sequence 8, Appli
17	1502	98.7	348	3	US-09-503-444A-8	Sequence 8, Appli
18	522	34.3	361	3	US-08-652-265-22	Sequence 22, Appl
19	522	34.3	361	3	US-08-834-497A-22	Sequence 22, Appl
20	522	34.3	361	3	US-09-503-444A-22	Sequence 22, Appl
21	516	33.9	364	4	US-08-914-372C-11	Sequence 11, Appl
22	513	33.7	365	3	US-08-652-265-23	Sequence 23, Appl
23	513	33.7	365	3	US-08-834-497A-23	Sequence 23, Appl
24	513	33.7	365	3	US-09-503-444A-23	Sequence 23, Appl
25	505	33.2	274	2	US-08-484-905-107	Sequence 107, App
26	505	33.2	274	3	US-08-481-985B-107	Sequence 107, App
27	505	33.2	274	3	US-08-370-476-107	Sequence 107, App
28	505	33.2	341	3	US-08-890-719-38	Sequence 38, Appl
29	504	33.1	365	2	US-08-484-905-97	Sequence 97, Appl
30	504	33.1	365	3	US-08-481-985B-97	Sequence 97, Appl
31	504	33.1	365	3	US-08-370-476-97	Sequence 97, Appl
32	503	33.0	274	2	US-08-484-905-108	Sequence 108, App
33	503	33.0	274	3	US-08-481-985B-108	Sequence 108, App
34	503	33.0	274	3	US-08-370-476-108	Sequence 108, App
35	503	33.0	365	2	US-08-484-905-100	Sequence 100, App
36	503	33.0	365	3	US-08-481-985B-100	Sequence 100, App
37	503	33.0	365	3	US-08-370-476-100	Sequence 100, App
38	502	33.0	274	1	US-08-222-851-1	Sequence 1, Appli
39	502	33.0	363	4	US-08-914-372C-37	Sequence 37, Appl
40	502	33.0	365	2	US-08-484-905-99	Sequence 99, Appl
41	502	33.0	365	3	US-08-481-985B-99	Sequence 99, Appl
42	502	33.0	365	3	US-08-370-476-99	Sequence 99, Appl
43	501	32.9	274	2	US-08-484-905-106	Sequence 106, App
44	501	32.9	274	3	US-08-481-985B-106	Sequence 106, App
45	501	32.9	274	3	US-08-370-476-106	Sequence 106, App

ALIGNMENTS

```
RESULT 1
US-09-094-964-1
; Sequence 1, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-1

  Query Match           100.0%;  Score 1522;  DB 4;  Length 276;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-142;
```

```
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-08-652-265-2
; Sequence 2, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-2

  Query Match           100.0%;  Score 1522;  DB 3;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
US-08-834-497A-2
; Sequence 2, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-2

  Query Match             100.0%;  Score 1522;  DB 3;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-09-503-444A-2
; Sequence 2, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-2

  Query Match             100.0%;  Score 1522;  DB 3;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 5
US-09-277-457-2
; Sequence 2, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-277-457-2

  Query Match             100.0%;  Score 1522;  DB 4;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202
```

```
Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 6
US-09-679-729-2
; Sequence 2, Application US/09679729
; Patent No. 6509442
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 24065-004 DIV
;  CURRENT APPLICATION NUMBER: US/09/679,729
;  CURRENT FILING DATE:  2000-10-04
;  PRIOR APPLICATION NUMBER: 09/277,457
;  PRIOR FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-679-729-2

  Query Match             100.0%;  Score 1522;  DB 4;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 3e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 7
US-09-094-964-2
; Sequence 2, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-2

  Query Match             99.4%;  Score 1513;  DB 4;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 8
US-08-652-265-6
; Sequence 6, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
```

```
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,265
;       FILING DATE:  23-MAY-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M.
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  17957-000500
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  348 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-652-265-6

  Query Match             99.4%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 9
US-08-834-497A-6
; Sequence 6, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-6

  Query Match             99.4%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 10
US-09-503-444A-6
; Sequence 6, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
```

```
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-6

  Query Match             99.4%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 11
US-08-652-265-4
; Sequence 4, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-4

  Query Match             99.3%;  Score 1511;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.7e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 12
US-08-834-497A-4
; Sequence 4, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
```

```
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-4

  Query Match             99.3%;  Score 1511;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.7e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 13
US-09-503-444A-4
; Sequence 4, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-4

  Query Match             99.3%;  Score 1511;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.7e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 14
US-09-094-964-3
; Sequence 3, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-3

  Query Match           98.7%;  Score 1502;  DB 4;  Length 276;
  Best Local Similarity  99.3%;  Pred. No. 2.1e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||| ||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 15
US-08-652-265-8
; Sequence 8, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-8

  Query Match           98.7%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity  99.3%;  Pred. No. 2.9e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                |||||||||||||||||| ||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

Search completed: May 4, 2004, 11:36:35
Job time : 16.3333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:32:18 ; Search time 13.6667 Seconds
(without alignments)
1942.600 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1140	75.0	359	2	JC5382	hereditary hemochr
2	542.5	35.7	341	2	A57136	class I histocompa
3	523	34.4	361	1	HLRB	MHC class I histoc
4	523	34.4	361	2	I46858	MHC class I RLA pr
5	520	34.2	332	2	S06424	MHC class I histoc
6	517	34.0	365	2	I36961	MHC class I protei
7	516	33.9	361	2	B27638	MHC class I histoc
8	515	33.9	365	2	I83063	A11.2 - human
9	514	33.8	365	2	A47636	MHC class I histoc
10	514	33.8	365	2	I56039	HLA-A30.3 precurso
11	512	33.7	370	1	HLHUA3	MHC class I histoc
12	510	33.6	365	2	I38439	MHC class I histoc
13	509	33.5	365	2	I37542	MHC class I histoc
14	509	33.5	365	2	I38442	gene HLA-A-0205 pr
15	509	33.5	365	2	I61902	MHC class I histoc
16	508	33.4	365	2	I72170	MHC class I histoc
17	508	33.4	365	2	I38441	gene HLA-A-6802 pr
18	506	33.3	365	1	HLHUA2	MHC class I histoc
19	506	33.3	365	2	I37482	MHC class I histoc
20	506	33.3	365	2	I38519	MHC class I histoc
21	506	33.3	365	2	I84448	MHC class I histoc
22	505	33.2	365	2	I38610	MHC class I histoc
23	505	33.2	365	2	I37470	HLA-A*0210 - human
24	504	33.2	355	2	T28149	MHC class I histoc
25	504	33.2	364	2	S03535	class I histocompa
26	503	33.1	365	2	I37476	MHC class I histoc
27	503	33.1	365	2	I37478	MHC class I histoc
28	503	33.1	365	2	I38443	gene HLA-A-0203 pr
29	503	33.1	365	2	I61857	MHC HLA-A2.4a chai
30	502.5	33.1	341	2	JC5663	major histocompati
31	502	33.0	357	2	I36965	MHC class I protei
32	501.5	33.0	362	2	A45845	MHC class I histoc
33	501	33.0	365	2	I61856	MHC class I histoc
34	501	33.0	365	2	I54493	MHC class I histoc
35	500	32.9	273	1	HLHU69	MHC class I histoc
36	500	32.9	365	2	S77963	MHC class I histoc
37	500	32.9	365	2	S01171	class I histocompa
38	500	32.9	365	2	I54416	HLA-AW24 protein -
39	499	32.8	365	2	I37483	HLA-Aw34.2 antigen
40	498	32.8	273	1	HLHUAW	MHC class I histoc
41	498	32.8	360	2	A27638	MHC class I histoc
42	498	32.8	365	2	I72171	HLA-Aw33.1, HLA-Aw
43	497.5	32.7	339	2	I56071	MHC class I histoc
44	497	32.7	279	2	JX0353	zinc-alpha 2-glyco
45	497	32.7	362	2	I68724	MHC class I histoc

ALIGNMENTS

RESULT 1
JC5382
hereditary hemochromatosis protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C;Accession: JC5382
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Biochem. Biophys. Res. Commun. 230, 35-39, 1997
A;Title: Identification of a mouse homolog for the human hereditary haemochromatosis ca
A;Reference number: JC5382; MUID:97148566; PMID:9020055
A;Accession: JC5382
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-359 <HAS>
A;Cross-references: GB:U66849; NID:g1519484; PIDN:AAB07525.1; PID:g1519485
C;Comment: This protein plays a role in iron metabolism.
C;Genetics:
A;Gene: mr2
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-359/Product: hereditary haemochromatosis protein #status predicted <MAT>
F;30-117/Domain: alpha 1 #status predicted <ALF1>
F;118-217/Domain: alpha 2 #status predicted <ALF2>
F;218-309/Domain: alpha 3 #status predicted <ALF3>
F;314-340/Domain: transmembrane #status predicted <TRM>
F;341-359/Domain: intracellular #status predicted <INT>

Query Match 75.0%; Score 1140; DB 2; Length 359;
Best Local Similarity 72.2%; Pred. No. 2.3e-87;
Matches 203; Conservative 30; Mismatches 40; Indels 8; Gaps 1;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| ||||||| |:|||||        ||| |||:||||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:||||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

RESULT 2
A57136
class I histocompatibility antigen related protein MR1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 23-Jul-1999
C;Accession: A57136
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Science 269, 693-695, 1995
A;Title: A gene outside the human MHC related to classical HLA class I genes.
A;Reference number: A57136; MUID:95350662; PMID:7624800
A;Accession: A57136
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-341 <HAS>
A;Cross-references: GB:U22963; NID:g940353; PIDN:AAC50174.1; PID:g940354
C;Genetics:
A;Gene: GDB:HLALS
A;Cross-references: GDB:683188; OMIM:600764
A;Map position: 1q25.3-1q25.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 35.7%; Score 542.5; DB 2; Length 341;
Best Local Similarity 39.5%; Pred. No. 1.2e-37;
Matches 107; Conservative 50; Mismatches 111; Indels 3; Gaps 3;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :   |   | || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| |||  |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291
```

RESULT 3
HLRB
MHC class I histocompatibility antigen RLA alpha chain precursor (RL-5) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 22-Jun-1999
C;Accession: A02193
R;Tykocinski, M.L.; Marche, P.N.; Max, E.E.; Kindt, T.J.
J. Immunol. 133, 2261-2269, 1984
A;Title: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expr
A;Reference number: A02193; MUID:84290724; PMID:6432910
A;Accession: A02193
A;Molecule type: mRNA
A;Residues: 1-361 <TYK>
A;Cross-references: GB:K02441; NID:g1293894; PIDN:AAA98729.1; PID:g165496
A;Note: the source of this protein is a T-lymphoid cell line (RL-5), which has been tran
C;Comment: In contrast to the many antigens expressed in mouse (K, D, and L) and human (
MHC may therefore differ from the HLA and H-2 loci in having limited complexity.
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: class I histocompatibility antigen RLA alpha chain #status predicted <
F;25-307/Domain: extracellular #status predicted <EXT>
F;25-114/Domain: alpha-1 <EX1>
F;115-206/Domain: alpha-2 <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;308-329/Domain: transmembrane #status predicted <TMM>
F;330-361/Domain: intracellular #status predicted <INT>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 34.4%; Score 523; DB 1; Length 361;
Best Local Similarity 40.1%; Pred. No. 5.3e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
                | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
               |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |     |||| || :||  |:: | :| :   :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 4
I46858
MHC class I RLA precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 21-Jan-2000
C;Accession: I46858
R;Marche, P.N.; Tykocinski, M.L.; Max, E.E.; Kindt, T.J.
Immunogenetics 21, 71-82, 1985
A;Title: Structure of a functional rabbit class I MHC gene: Similarity to human class I
A;Reference number: I46858; MUID:85103547; PMID:3917974
A;Accession: I46858
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-361 <MAR>
A;Cross-references: GB:K02819; NID:g165497; PIDN:AAA98730.1; PID:g165498
C;Genetics:
A;Introns: 25/1; 115/1; 207/1; 299/1; 337/1; 348/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 34.4%; Score 523; DB 2; Length 361;
Best Local Similarity 40.1%; Pred. No. 5.3e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
                | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
               |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |     |||| || :||  |:: | :| :   :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || ||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 5
S06424
MHC class I histocompatibility antigen Ch25 alpha chain precursor - chimpanzee
N;Alternate names: MHC Ch1A chain
C;Species: Pan troglodytes (chimpanzee)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C;Accession: S06424; I36959
R;Lawlor, D.A.; Ward, F.E.; Ennis, P.D.; Jackson, A.P.; Parham, P.
Nature 335, 268-271, 1988
A;Title: HLA-A and B polymorphisms predate the divergence of humans and chimpanzees.
A;Reference number: S06424; MUID:88319000; PMID:3412487
A;Accession: S06424
A;Molecule type: mRNA
A;Residues: 1-332 <LAW>
R;Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.
J. Immunol. 142, 3937-3950, 1989
A;Title: Diversity and diversification of HLA-A,B,C alleles.
A;Reference number: I36956; MUID:89235215; PMID:2715640
A;Accession: I36959
A;Molecule type: mRNA
A;Residues: 1-332 <RES>
A;Cross-references: GB:M24047; NID:g176818; PIDN:AAA35426.1; PID:g553155
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-114/Domain: alpha-1 #status predicted <EX1>
F;115-206/Domain: alpha-2 #status predicted <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;307-331/Domain: transmembrane #status predicted <TMM>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 34.2%; Score 520; DB 2; Length 332;
Best Local Similarity 40.1%; Pred. No. 8.5e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |    | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::| |        ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 6
I36961
MHC class I protein - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I36961
R;Lawlor, D.A.; Warren, E.; Ward, F.E.; Parham, P.
Immunol. Rev. 113, 147-185, 1990
A;Title: Comparison of class I MHC alleles in humans and apes.
A;Reference number: I36961; MUID:90201944; PMID:1690682
A;Accession: I36961
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30678; NID:g176822; PIDN:AAA87970.1; PID:g176823
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 34.0%; Score 517; DB 2; Length 365;
Best Local Similarity 39.7%; Pred. No. 1.7e-35;
Matches 110; Conservative 45; Mismatches 114; Indels 8; Gaps 7;

```
Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |    | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::| |        ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 7
B27638
MHC class I histocompatibility antigen alpha chain precursor (BL3-7) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 16-Feb-1997
C;Accession: B27638
R;Ennis, P.D.; Jackson, A.P.; Parham, P.
J. Immunol. 141, 642-651, 1988
A;Title: Molecular cloning of bovine class I MHC cDNA.
A;Reference number: A92826; MUID:88258075; PMID:3133413
A;Accession: B27638
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-361 <ENN>
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: MHC class I histocompatibility antigen, BoLA alpha chain (BL3-7) #sta
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.9%; Score 516; DB 2; Length 361;
Best Local Similarity 38.9%; Pred. No. 2e-35;
Matches 109; Conservative 50; Mismatches 113; Indels 8; Gaps 7;

```
Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISS 59
              |  |||| | : | |    ||    | |:|||||  || :| :  : | ||| ||:
Db         23 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 81

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |   :| ||  |:   :| |:  || :| : ||::  |     |:|::|
Db         82 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 141

Qy        118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        142 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 200

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              | :  ||   |||| ::    |||| || :||: |:: | :: :    :: |  :  |
Db        201 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 259

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:| |  | || ||||||||:|:| || :|| : ||
Db        260 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 299
```

RESULT 8

```
I83063
A11.2 - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I83063
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.;
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I83063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16842; NID:g540517; PIDN:BAA04118.1; PID:g487911
C;Genetics:
A;Gene: A1102
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.9%;  Score 515;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGKPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|       ||  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
             :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :  ||   :||| ::    |||| || :|| ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
             ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 9
A47636
MHC class I histocompatibility antigen HLA-A11 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 28-Apr-1995 #text_change 23-Jul-1999
C;Accession: S03536; S03694; A47636; I60129
R;Mayer, W.E.; Jonker, M.; Klein, D.; Ivanyi, P.; van Seventer, G.; Klein, J.
EMBO J. 7, 2765-2774, 1988
A;Title: Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-spec
A;Reference number: S01171; MUID:89030641; PMID:2460344
A;Accession: S03536
A;Molecule type: mRNA
A;Residues: 1-365 <MAY>
A;Cross-references: EMBL:X13111; NID:g32138; PIDN:CAA31503.1; PID:g32139
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
A;Accession: S03694
A;Molecule type: mRNA
A;Residues: 1-42,'K',44-298 <MA2>
A;Cross-references: EMBL:X13112; NID:g32142; PIDN:CAA31504.1; PID:g32143
A;Note: this allele is designated A*1102 (formerly A11K, A11.2)
R;Cowan, E.P.; Jelachich, M.L.; Biddison, W.E.; Coligan, J.E.
Immunogenetics 25, 241-250, 1987
A;Title: DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows identification
A;Reference number: A47636; MUID:87192928; PMID:2437024
A;Accession: A47636
A;Molecule type: DNA
A;Residues: 26-365 <COW>
A;Cross-references: GB:M16007; GB:M16008; GB:M16009; GB:M16010; NID:g184130; PIDN:AAA654
A;Note: the authors translated the codon GAC for residue 89 as Ala, CCG for residue 104
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I60129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16841; NID:g540516; PIDN:BAA04117.1; PID:g487909
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: class I histocompatibility antigen alpha chain #status predicted <MAT
F;25-298/Domain: extracellular #status predicted <EXT>
F;220-285/Domain: immunoglobulin homology <IMM>
F;299-337/Domain: transmembrane #status predicted <TMM>
F;338-365/Domain: intracellular #status predicted <INT>

  Query Match             33.8%;  Score 514;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 3e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|       ||  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
             :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :  ||   :||| ::    |||| || :|| ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
             ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 10
I56039
HLA-A30.3 precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I56039
R;Kato, K.; Trapani, J.A.; Allopenna, J.; Dupont, B.; Yang, S.Y.
J. Immunol. 143, 3371-3378, 1989
A;Title: Molecular analysis of the serologically defined HLA-Aw19 antigens. A genetical
A;Reference number: I56039; MUID:90038496; PMID:2478623
A;Accession: I56039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30576; NID:g187646; PIDN:AAA59612.1; PID:g386878
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.8%;  Score 514;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 3e-35;
  Matches  109;  Conservative   48;  Mismatches  112;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
```

```
Db         26 SHSMRYFFTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :: |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYEQHAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  |: :||  :  | | ||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITQRKWEAAR-WAEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HLHUA3
MHC class I histocompatibility antigen HLA-A3 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 02-Sep-1997
C;Accession: A02192
R;Strachan, T.; Sodoyer, R.; Damotte, M.; Jordan, B.R.
EMBO J. 3, 887-894, 1984
A;Title: Complete nucleotide sequence of a functional class I HLA gene, HLA-A3: implicat
A;Reference number: A02192; MUID:84207948; PMID:6609814
A;Accession: A02192
A;Molecule type: DNA
A;Residues: 1-370 <STR>
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
A;Introns: 30/1; 120/1; 212/1; 304/1; 343/1; 354/1; 370/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-370/Product: class I histocompatibility antigen HLA-A3 alpha chain #status predicte
F;30-312/Domain: extracellular #status predicted <EXT>
F;30-119/Domain: alpha-1 <EX1>
F;120-211/Domain: alpha-2 <EX2>
F;225-290/Domain: immunoglobulin homology <IMM>
F;313-337/Domain: transmembrane #status predicted <TMM>
F;338-370/Domain: intracellular #status predicted <INT>
F;115/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;232-288/Disulfide bonds: #status predicted

  Query Match             33.7%;  Score 512;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 4.5e-35;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | ::::|        ||  |:   :| |:  |||:|:: ||::  |      || :  |||
Db         90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db        150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 12
I38439
MHC class I histocompatibility antigen HLA-A*8001 precursor - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 21-Jan-2000
C;Accession: I59638; I38439
R;Domena, J.D.; Hildebrand, W.H.; Bias, W.B.; Parham, P.
Tissue Antigens 42, 156-159, 1993
A;Title: A sixth family of HLA-A alleles defined by HLA-A*8001.
A;Reference number: I59638; MUID:94112691; PMID:8284791
A;Accession: I59638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <DOM>
A;Cross-references: GB:L18898; NID:g306853; PIDN:AAA17012.1; PID:g306854
R;Balas, A.; Garcia-Sanchez, F.; Gomez-Reino, F.; Vicario, J.L.
Immunogenetics 39, 452, 1994
A;Title: Characterization of a new and highly distinguishable HLA-A allele in a Spanish
A;Reference number: I38439; MUID:94245293; PMID:8188325
A;Accession: I38439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <BAL>
A;Cross-references: EMBL:U03754; NID:g432407; PIDN:AAC04322.1; PID:g432408
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.6%;  Score 510;  DB 2;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 6.5e-35;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              : ::::|        :  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 13
I37542
MHC class I histocompatibility antigen HLA-A2 alpha chain (allele A*0216) precursor - h
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I37542; S49582
R;Barouch, D.; Krausa, P.; Bodmer, J.; Browning, M.J.; McMichael, A.J.
Immunogenetics 41, 388, 1995
A;Title: Identification of a novel HLA-A2 subtype, HLA-A*0216.
A;Reference number: I37542; MUID:95278976; PMID:7759139
A;Accession: I37542
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:Z46633; NID:g575248; PIDN:CAA86602.1; PID:g575249
```

```
A;Note: submitted to the EMBL Data Library, November 1994
C;Genetics:
A;Gene: hla-A
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-35;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              ||| | | |     |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
               :: :|       ||  |:   :| |:  |||:| : ||:: |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  :|   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGECVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||  |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
I38442
gene HLA-A-0205 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I38442
R;Holmes, N.; Ennis, P.; Wan, A.M.; Denney, D.W.; Parham, P.
J. Immunol. 139, 936-941, 1987
A;Title: Multiple genetic mechanisms have contributed to the generation of the HLA-A2/A2
A;Reference number: I38441; MUID:87252273; PMID:3496393
A;Accession: I38442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:U03862; NID:g432436; PIDN:AAA03603.1; PID:g432437
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 7.9e-35;
  Matches  110;  Conservative   44;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              ||| | :   |     |   | |:|||||  || :| :  |||:||| ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
               :: :|       ||  |:   :| |:  |||:| : ||:: |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||  |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
I61902
MHC class I histocompatibility antigen HLA-A alpha chain precursor - human (isolate A*0
C;Species: Homo sapiens (man)
A;Variety: isolate A*0212
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C;Accession: I61902
R;Belich, M.P.; Madrigal, J.A.; Hildebrand, W.H.; Zemmour, J.; Williams, R.C.; Luz, R.;
Nature 357, 326-329, 1992
A;Title: Unusual HLA-B alleles in two tribes of Brazilian Indians.
A;Reference number: I37120; MUID:92269955; PMID:1317015
A;Accession: I61902
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M84378; NID:g187625; PIDN:AAA59604.1; PID:g187626
A;Experimental source: cell line KRC 033; isolate A*0212
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: MHC class I histocompatibility antigen HLA-A alpha chain #status pred
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.5%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-35;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              ||| | | |     |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
               :: :|       ||  |:   :| |:  |||:| : ||:: |     || :| |||
Db         85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db        145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|||||  |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: May  4, 2004, 11:39:25
Job time : 13.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:32:18 ; Search time 8.33333 Seconds
(without alignments)
1724.564 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1513	99.5	348	1	HFE_HUMAN	Q30201 homo sapien
2	1513	99.5	348	1	HFE_PANTR	P60018 pan troglod
3	1238	81.4	348	1	HFE_DICSU	Q9gl42 dicerorhinu
4	1236	81.3	348	1	HFE_CERSI	Q9gkz0 ceratotheri
5	1232	81.1	348	1	HFE_RHIUN	Q9gl41 rhinoceros
6	1229	80.9	348	1	HFE_DICBI	Q9gl43 diceros bic
7	1156	76.1	360	1	HFE_RAT	O35799 rattus norv
8	1140	75.0	359	1	HFE_MOUSE	P70387 mus musculu
9	523	34.4	361	1	HA1A_RABIT	P01894 oryctolagus
10	523	34.4	361	1	HA1B_RABIT	P06140 oryctolagus
11	517	34.0	365	1	1A01_PANTR	P16209 pan troglod
12	516	33.9	364	1	HA1B_BOVIN	P13753 bos taurus
13	514	33.8	365	1	1A11_HUMAN	P13746 homo sapien
14	512	33.7	365	1	1A03_HUMAN	P04439 homo sapien
15	510	33.6	365	1	1A80_HUMAN	Q09160 homo sapien
16	508	33.4	365	1	1A31_HUMAN	P16189 homo sapien
17	506	33.3	365	1	1A02_HUMAN	P01892 homo sapien
18	506	33.3	365	1	1A30_HUMAN	P16188 homo sapien
19	506	33.3	365	1	1A74_HUMAN	P30459 homo sapien
20	504	33.2	365	1	1A03_PANTR	P13748 pan troglod
21	503	33.1	365	1	1A33_HUMAN	P16190 homo sapien
22	503	33.1	365	1	1A36_HUMAN	P30455 homo sapien
23	503	33.1	365	1	1A68_HUMAN	P01891 homo sapien
24	501.5	33.0	362	1	HA19_CANFA	P18466 canis famil
25	501	33.0	365	1	1A01_HUMAN	P30443 homo sapien
26	500	32.9	273	1	1A69_HUMAN	P10316 homo sapien
27	500	32.9	365	1	1A04_PANTR	P13749 pan troglod
28	500	32.9	365	1	1A24_HUMAN	P05534 homo sapien
29	498	32.8	360	1	HA1A_BOVIN	P13752 bos taurus
30	497	32.7	296	1	ZA2G_RAT	Q63678 rattus norv
31	497	32.7	362	1	1B47_HUMAN	P30485 homo sapien
32	496	32.6	365	1	1A23_HUMAN	P30447 homo sapien
33	493	32.4	363	1	1B04_GORGO	P30382 gorilla gor
34	492	32.4	295	1	ZA2G_HUMAN	P25311 homo sapien
35	492	32.4	322	1	HA10_MOUSE	P01898 mus musculu
36	492	32.4	362	1	1B37_HUMAN	P18463 homo sapien
37	492	32.4	371	1	HA12_RAT	P16391 rattus norv
38	491	32.3	365	1	1A34_HUMAN	P30453 homo sapien
39	491	32.3	365	1	1A66_HUMAN	P30457 homo sapien
40	490	32.2	338	1	HLAG_HUMAN	P17693 homo sapien
41	490	32.2	362	1	1B27_HUMAN	P03989 homo sapien
42	490	32.2	366	1	1C02_GORGO	P30385 gorilla gor
43	490	32.2	366	1	1C04_GORGO	P30387 gorilla gor
44	489	32.2	359	1	1B01_PANTR	P13750 pan troglod
45	489	32.2	365	1	1A01_PONPY	P16211 pongo pygma

ALIGNMENTS

RESULT 1
HFE_HUMAN
ID HFE_HUMAN STANDARD; PRT; 348 AA.
AC Q30201; O75929; O75930; O75931; Q96KU5; Q96KU7; Q96KU8; Q9HC64;
AC Q9HC68; Q9HC70; Q9HC83;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hereditary hemochromatosis protein precursor (HLA-H).
GN HFE OR HLAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=96331279; PubMed=8696333;
RA Feder J.N., Gnirke A., Thomas W., Tsuchihashi Z., Ruddy D.A.,
RA Basava A., Dormishian F., Domingo R., Ellis M.C. Jr., Fullan A.,
RA Hinton L.M., Jones N.L., Kimmel B.E., Kronmal G.S., Lauer P.,
RA Lee V.K., Loeb D.B., Mapa F.A., McClelland E., Meyer N.C.,
RA Mintier G.A., Moeller N., Moore T., Morikang E., Prass C.E.,
RA Quintana L., Starnes S.M., Schatzman R.C., Brunke K.J.,
RA Drayna D.T., Risch N.J., Bacon B.R., Wolff R.K.;
RT "A novel MHC class I-like gene is mutated in patients with hereditary
RT haemochromatosis.";
RL Nat. Genet. 13:399-409(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Albig W., Burmester N., Bode C., Doenecke D., Drabent B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97294057; PubMed=9149941;
RA Ruddy D.A., Kronmal G.S., Lee V.K., Mintier G.A., Quintana L.,
RA Domingo R. Jr., Meyer N.C., Irrinki A., McClelland E.E., Fullan A.,
RA Mapa F.A., Moore T., Thomas W., Loeb D.B., Harmon C., Tsuchihashi Z.,
RA Wolff R.K., Schatzman R.C., Feder J.N.;
RT "A 1.1-Mb transcript map of the hereditary hemochromatosis locus.";
RL Genome Res. 7:441-456(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gasparini P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).
RX MEDLINE=99180629; PubMed=10079302;
RA Rhodes D.A., Trowsdale J.;
RT "Alternate splice variants of the hemochromatosis gene Hfe.";
RL Immunogenetics 49:357-359(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 5; 6 AND 7).
RA Oliva R., Sanchez M.;
RT "Identification of different alternative splicing forms of the HFE
RT gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.

RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 7; 8; 9 AND 10).
RX MEDLINE=20448010; PubMed=11001625;
RA Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.-Y.,
RA David V., Mosser J.;
RT "The HFE gene undergoes alternate splicing processes.";
RL Blood Cells Mol. Dis. 26:155-162(2000).
RN [8]
RP FUNCTION.
RX MEDLINE=98132614; PubMed=9465039;
RA Feder J.N., Penny D.M., Irrinki A., Lee V.K., Lebron J.A., Watson N.,
RA Tsuchihashi Z., Sigal E., Bjorkman P.J., Schatzman R.C.;
RT "The hemochromatosis gene product complexes with the transferrin
RT receptor and lowers its affinity for ligand binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1472-1477(1998).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=98206473; PubMed=9546397;
RA Lebron J.A., Bennett M.J., Vaughn D.E., Chirino A.J., Snow P.M.,
RA Mintier G.A., Feder J.N., Bjorkman P.J.;
RT "Crystal structure of the hemochromatosis protein HFE and
RT characterization of its interaction with transferrin receptor.";
RL Cell 93:111-123(1998).
RN [10]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=97260408; PubMed=9106528;
RA Carella M., D'Ambrosio L., Totaro A., Grifa A., Valentino M.A.,
RA Piperno A., Girelli D., Roetto A., Franco B., Gasparini P.,
RA Camaschella C.;
RT "Mutation analysis of the HLA-H gene in Italian hemochromatosis
RT patients.";
RL Am. J. Hum. Genet. 60:828-832(1997).
RN [11]
RP VARIANT HH/PCT TYR-282.
RX MEDLINE=97176837; PubMed=9024376;
RA Roberts A.G., Whatley S.D., Morgan R.R., Worwood M., Elder G.H.;
RT "Increased frequency of the haemochromatosis Cys282Tyr mutation in
RT sporadic porphyria cutanea tarda.";
RL Lancet 349:321-323(1997).
RN [12]
RP VARIANT HH/PCT ASP-63.
RX MEDLINE=98085904; PubMed=9425935;
RA Sampietro M., Piperno A., Lupica L., Arosio C., Vergani A.,
RA Corbetta N., Malosio I., Mattioli M., Fracanzani A.L.,
RA Cappellini M.D., Fiorelli G., Fargion S.;
RT "High prevalence of the His63Asp HFE mutation in Italian patients with
RT porphyria cutanea tarda.";
RL Hepatology 27:181-184(1998).
RN [13]
RP VARIANTS HH/PCT ASP-63 AND TYR-282.
RX MEDLINE=98281650; PubMed=9620340;
RA Bonkovsky H.L., Poh-Fitzpatrick M., Pimstone N., Obando J.,
RA Di Bisceglie A., Tattrie C., Tortorelli K., LeClair P., Mercurio M.G.,
RA Lambrecht R.W.;
RT "Porphyria cutanea tarda, hepatitis C, and HFE gene mutations in North
RT America.";
RL Hepatology 27:1661-1669(1998).
RN [14]
RP VARIANTS HH ASP-63; CYS-65 AND TYR-282.
RX MEDLINE=99211934; PubMed=10194428;
RA Mura C., Raguenes O., Ferec C.;
RT "HFE mutations analysis in 711 hemochromatosis probands: evidence for
RT S65C implication in mild form of hemochromatosis.";
RL Blood 93:2502-2505(1999).
RN [15]
RP VARIANTS HH CYS-65; ARG-93 AND THR-105.
RX MEDLINE=20042794; PubMed=10575540;
RA Barton J.C., Sawada-Hirai R., Rothenberg B.E., Acton R.T.;
RT "Two novel missense mutations of the HFE gene (I105T and G93R) and
RT identification of the S65C mutation in Alabama hemochromatosis
RT probands.";
RL Blood Cells Mol. Dis. 25:147-155(1999).
RN [16]
RP VARIANTS VP ASP-63 AND HIS-127, VARIANT HH MET-330, AND VARIANTS
RP MET-53 AND MET-59.
RX MEDLINE=99330560; PubMed=10401000;
RA de Villiers J.N.P., Hillermann R., Loubser L., Kotze M.J.;
RT "Spectrum of mutations in the HFE gene implicated in haemochromatosis
RT and porphyria.";
RL Hum. Mol. Genet. 8:1517-1522(1999).
RN [17]
RP VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=99140260; PubMed=10094552;
RA Merryweather-Clarke A.T., Simonsen H., Shearman J.D., Pointon J.J.,
RA Norgaard-Pedersen B., Robson K.J.H.;
RT "A retrospective anonymous pilot study in screening newborns for HFE
RT mutations in Scandinavian populations.";
RL Hum. Mutat. 13:154-159(1999).
RN [18]
RP VARIANT HH CYS-65.
RA Fagan E., Payne S.J.;
RT "A novel missense mutation S65C in the HFE gene with a possible role
RT in hereditary haemochromatosis.";
RL Hum. Mutat. 13:507-508(1999).
RN [19]
RP VARIANT LYS-277.
RX MEDLINE=20081073; PubMed=10612845;
RA Bradbury R., Fagan E., Payne S.J.;
RT "Two novel polymorphisms (E277K and V212V) in the haemochromatosis
RT gene HFE.";
RL Hum. Mutat. 15:120-120(2000).
CC -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC affinity for iron-loaded transferrin.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q30201-1; Sequence=Displayed;
CC Name=2; Synonyms=delE2;
CC IsoId=Q30201-2; Sequence=VSP_003218;
CC Name=3; Synonyms=del14E4;
CC IsoId=Q30201-3; Sequence=VSP_003225;
CC Name=4; Synonyms=delE214E4;
CC IsoId=Q30201-4; Sequence=VSP_003218, VSP_003225;
CC Name=5;
CC IsoId=Q30201-5; Sequence=VSP_003219;
CC Name=6;
CC IsoId=Q30201-6; Sequence=VSP_003220;
CC Name=7; Synonyms=delE3;
CC IsoId=Q30201-7; Sequence=VSP_003221;
CC Name=8; Synonyms=1043-2283del,intron6ins;
CC IsoId=Q30201-8; Sequence=VSP_003226, VSP_003227;
CC Name=9; Synonyms=delE3-7;
CC IsoId=Q30201-9; Sequence=VSP_003223, VSP_003224;
CC Name=10; Synonyms=562-878del;
CC IsoId=Q30201-10; Sequence=VSP_003222;
CC -!- TISSUE SPECIFICITY: In all tissues tested except brain.
CC -!- DISEASE: Defects in HFE are a cause of hereditary hemochromatosis
CC (HH) [MIM:235200]. HH is an autosomal recessive inborn disorder of
CC iron metabolism, frequent among caucasians. HH is characterized by
CC abnormal intestinal iron absorption and progressive increase of
CC total body iron, which results in midlife in clinical
CC complications including cirrhosis, cardiopathy, diabetes,
CC endocrine dysfunctions, arthropathy, and susceptibility to liver
CC cancer. Since the disease complications can be effectively
CC prevented by regular phlebotomies, early diagnosis is most
CC important to provide a normal life expectancy to the affected
CC subjects.
CC -!- DISEASE: Defects in HFE are a cause of porphyria cutanea tarda
CC (PCT), a disorder characterized by light-sensitive dermatitis and
CC presence of large amounts of uroporphyrin in urine. Iron overload
CC is often present in association with varying degrees of liver
CC damage. PCT is the most common form of porphyria worldwide. It
CC occurs in two forms: the sporadic type (PCT type I) and the
CC familial type (PCT type II), which is inherited in an autosomal

```
  Query Match           99.5%;  Score 1513;  DB 1;  Length 348;
  Best Local Similarity  99.6%;  Pred. No. 1.9e-118;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 2
HFE_PANTR
ID   HFE_PANTR      STANDARD;      PRT;   348 AA.
AC   P60018;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Hereditary hemochromatosis protein precursor (HLA-H).
GN   HFE OR HLAH.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22184165; PubMed=12196404;
RA   Toomajian C., Kreitman M.;
RT   "Sequence variation and haplotype structure at the Human HFE Locus.";
RL   Genetics 161:1609-1623(2002).
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF447807; AAN09793.1; -.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Transport; Iron transport; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  40108 MW;  432EB9A314A55BEA CRC64;

  Query Match           99.5%;  Score 1513;  DB 1;  Length 348;
  Best Local Similarity  99.6%;  Pred. No. 1.9e-118;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
HFE_DICSU
ID   HFE_DICSU      STANDARD;      PRT;   348 AA.
AC   Q9GL42;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Dicerorhinus sumatrensis (Sumatran rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
OX   NCBI_TaxID=89632;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007543; AAG23703.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
```

```
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     22       BY SIMILARITY.
FT   CHAIN       23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN      23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN     115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN     206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN     298    306       CONNECTING PEPTIDE.
FT   TRANSMEM   307    330       POTENTIAL.
FT   DOMAIN     331    348       CYTOPLASMIC TAIL.
FT   DISULFID   124    187       BY SIMILARITY.
FT   DISULFID   225    282       BY SIMILARITY.
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39740 MW;  518BFD357AB83B90 CRC64;

  Query Match             81.4%;  Score 1238;  DB 1;  Length 348;
  Best Local Similarity   81.3%;  Pred. No. 1.3e-95;
  Matches  222;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||| |||||||:|||||||||||||||:|||||  |:|||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||||  |:||||||:|||:|||||||:||||::|:|  |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: |||||||||||||||||||||||   ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 4
HFE_CERSI
ID   HFE_CERSI      STANDARD;      PRT;   348 AA.
AC   Q9GKZ0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX   NCBI_TaxID=9807;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007541; AAG23701.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     22       BY SIMILARITY.
FT   CHAIN       23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN      23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN     115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN     206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN     298    306       CONNECTING PEPTIDE.
FT   TRANSMEM   307    330       POTENTIAL.
FT   DOMAIN     331    348       CYTOPLASMIC TAIL.
FT   DISULFID   124    187       BY SIMILARITY.
FT   DISULFID   225    282       BY SIMILARITY.
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39822 MW;  2523016ECE9FBE91 CRC64;

  Query Match             81.3%;  Score 1236;  DB 1;  Length 348;
  Best Local Similarity   81.7%;  Pred. No. 1.9e-95;
  Matches  223;  Conservative   18;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |  ||     ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||| |||||||:|||||||||||||||:|||||  |:|||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||||  |:||||||:|||:|||||||:||||::|:|  |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |  |||||||||||||||||||||||   ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 5
HFE_RHIUN
ID   HFE_RHIUN      STANDARD;      PRT;   348 AA.
AC   Q9GL41;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Rhinoceros unicornis (Greater Indian rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
OX   NCBI_TaxID=9809;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
```

```
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007544; AAG23704.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39743 MW;  F2723D57A327A6B4 CRC64;

  Query Match             81.1%;  Score 1232;  DB 1;  Length 348;
  Best Local Similarity   81.0%;  Pred. No. 4e-95;
  Matches  221;  Conservative   20;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |  ||      ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              |||||||||||| |||||||:||||||||||||||||:|||||| |:| ||||||||
Db         86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:|||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: |||||||||||||||||||||||    ||
Db        266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 6
HFE_DICBI
ID   HFE_DICBI      STANDARD;      PRT;   348 AA.
AC   Q9GL43; Q9GK81;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein precursor.
GN   HFE.
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY007542; AAG23702.1; -.
DR   EMBL; AF301592; AAG39940.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     22       BY SIMILARITY.
FT   CHAIN        23    348       HEREDITARY HEMOCHROMATOSIS PROTEIN.
FT   DOMAIN       23    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    205       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      206    297       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      298    306       CONNECTING PEPTIDE.
FT   TRANSMEM    307    330       POTENTIAL.
FT   DOMAIN      331    348       CYTOPLASMIC TAIL.
FT   DISULFID    124    187       BY SIMILARITY.
FT   DISULFID    225    282       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   348 AA;  39818 MW;  4D95E7B01E48FB90 CRC64;

  Query Match             80.9%;  Score 1229;  DB 1;  Length 348;
  Best Local Similarity   81.0%;  Pred. No. 7.2e-95;
  Matches  221;  Conservative   20;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||:| || ||||||||||:||| |: |||| | |  ||      ||:||
Db         26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              ||:|||||||||| |||||||:|||||||||||||||||||:|||||| |:|||||||||
Db         86 QLTQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db        146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy        184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
              |||||||||| |:|||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db        206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265
```

```
Qy        244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |  ||||||||||||||||||||||||||    ||
Db        266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 7
HFE_RAT
ID   HFE_RAT        STANDARD;      PRT;   360 AA.
AC   O35799; O35175;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor (RT1-CAFE).
GN   HFE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Banasch M.W., Schaefer H., Schmidt W.E.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 42-303 FROM N.A.
RC   TISSUE=Small intestine;
RA   Sawada-Hirai R., Rothenberg B.E.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001517; CAA04799.1; -.
DR   EMBL; AF008587; AAB86597.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    360       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       26    127       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      128    218       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      219    310       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      311    319       CONNECTING PEPTIDE.
FT   TRANSMEM    320    340       POTENTIAL.
FT   DOMAIN      341    360       CYTOPLASMIC TAIL.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    238    295       BY SIMILARITY.
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    198    198       R -> K (IN REF. 2).
SQ   SEQUENCE   360 AA;  40988 MW;  CC819834EE3240B3 CRC64;

  Query Match             76.1%;  Score 1156;  DB 1;  Length 360;
  Best Local Similarity   73.6%;  Pred. No. 8.6e-89;
  Matches  206;  Conservative   29;  Mismatches   37;  Indels    8;  Gaps    1;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQ 64
              |||| ||||||: ||||  ||||||||||||| |: |||| ||| ||:  : |||:|||
Db         32 SHSLRYLFMGASKPDLGLPFFEALGYVDDQLFVSYNHESRRAEPRAPWILGQTSSQLWLQ 91

Qy         65 LSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWKY 116
              |||||||||:|| |||||| |:|||         ||| ||||||||: ||||| |:|||
Db         92 LSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLRVVPESHILQVILGCEVHEDNSTSGFWKY 151

Qy        117 GYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELG 176
              |||||||||||| ||:| ||||||| ||:||| |:|||||:| ||:|||| ||:|:|||
Db        152 GYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHRIRARQSRDYLQRDCPQQLKQVLELQ 211

Qy        177 RGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              |||| |||| ||||| |  |: |:|||:|||::||||||:|||| ||:|||:  |::|||
Db        212 RGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQNITMRWLKDSQPLDAKDVNPENVLP 271

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||:|||| |||| |::||||||||||||    ||
Db        272 NGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPLTATWE 311


RESULT 8
HFE_MOUSE
ID   HFE_MOUSE      STANDARD;      PRT;   359 AA.
AC   P70387;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor.
GN   HFE OR MR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98060831; PubMed=9396865;
RA   Riegert P., Gilfillan S., Nanda I., Schmid M., Bahram S.;
RT   "The mouse HFE gene.";
RL   Immunogenetics 47:174-177(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Lung;
RA   Hashimoto K.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 37-211 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=97148566; PubMed=9020055;
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "Identification of a mouse homolog for the human hereditary
RT   haemochromatosis candidate gene.";
RL   Biochem. Biophys. Res. Commun. 230:35-39(1997).
RN   [4]
RP   SEQUENCE OF 79-359 FROM N.A.
RC   STRAIN=129;
RA   Albig W., Drabent B., Doenecke D.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds to transferrin receptor (TFR) and reduces its
CC       affinity for iron-loaded transferrin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF007558; AAC03447.1; -.
DR   EMBL; U66849; AAB07525.1; -.
DR   EMBL; Y12650; CAA73197.1; -.
DR   EMBL; U80604; AAB51504.1; -.
DR   PIR; JC5382; JC5382.
DR   HSSP; Q30201; 1A6Z.
DR   MGD; MGI:109191; Hfe.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    359       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       25    126       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      127    217       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      218    309       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      310    318       CONNECTING PEPTIDE.
FT   TRANSMEM    319    339       POTENTIAL.
FT   DOMAIN      340    359       CYTOPLASMIC TAIL.
FT   DISULFID    136    199       BY SIMILARITY.
FT   DISULFID    237    294       BY SIMILARITY.
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   359 AA;  40548 MW;  4BDE6C27F9FF20B4 CRC64;

  Query Match             75.0%;  Score 1140;  DB 1;  Length 359;
  Best Local Similarity   72.2%;  Pred. No. 1.8e-87;
  Matches  203;  Conservative   30;  Mismatches   40;  Indels    8;  Gaps    1;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| |||||||| |: |||| ||| ||:  : |||:||
Db         30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
               |||||||||:|| ||||||| |:|||         ||| |||:|||: ||||| |:|:
Db         90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||||| ||:| |||| || ||:||: ||||:||| |||:||| ||::||||
Db        150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db        210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| |||||:|||| ||:| |:|||||||||||||   ||
Db        270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310


RESULT 9
HA1A_RABIT
ID   HA1A_RABIT     STANDARD;      PRT;   361 AA.
AC   P01894;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 11/11 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84290724; PubMed=6432910;
RA   Tykocinski M.L., Marche P.N., Max E.E., Kindt T.J.;
RT   "Rabbit class I MHC genes: cDNA clones define full-length transcripts
RT   of an expressed gene and a putative pseudogene.";
RL   J. Immunol. 133:2261-2269(1984).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to the
CC       immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; K02441; AAA98729.1; -.
DR   PIR; A02193; HLRB.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                ALPHA CHAIN 11/11.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40447 MW;  580B673323C1AE35 CRC64;

  Query Match             34.4%;  Score 523;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 3.1e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |   ||   |  :|||||  || :|  :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::   |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |    |||| || :||  |:: | :|  :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262
```

```
Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 10
HA1B_RABIT
ID   HA1B_RABIT     STANDARD;      PRT;   361 AA.
AC   P06140;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 19-1 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85103547; PubMed=3917974;
RA   Marche P.N., Tykocinski M.L., Max E.E., Kindt T.J.;
RT   "Structure of a functional rabbit class I MHC gene: similarity to
RT   human class I genes.";
RL   Immunogenetics 21:71-82(1985).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to the
CC       immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02819; AAA98730.1; -.
DR   PIR; I46858; I46858.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 19-1.
FT   DOMAIN        25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    329
FT   DOMAIN       330    361       CYTOPLASMIC.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40455 MW;  C06FBD8B87ED0546 CRC64;

  Query Match             34.4%;  Score 523;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 3.1e-36;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |  ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db         26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
              | :|  |    | |:  | :  :| |   ||| | : |||:  |    || :| |||
Db         85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db        145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy        181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   |||| |     |||| || :||  |:: | :| :    :: |  :  | ||
Db        204 QRADPPKAHVTHHPISDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||||:|:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 11
1A01_PANTR
ID   1A01_PANTR     STANDARD;      PRT;   365 AA.
AC   P16209;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   CHLA class I histocompatibility antigen, A-2 alpha chain precursor.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90201944; PubMed=1690682;
RA   Lawlor D.A., Warren E., Ward F.E., Parham P.;
RT   "Comparison of class I MHC alleles in humans and apes.";
RL   Immunol. Rev. 113:147-185(1990).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to the
CC       immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30678; AAA87970.1; -.
DR   PIR; I36961; I36961.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-2 ALPHA CHAIN.
FT   DOMAIN        25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
```

```
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   365 AA;  40848 MW;  FC452786BD038D3E CRC64;

  Query Match             34.0%;  Score 517;  DB 1;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 1e-35;
  Matches  110;  Conservative   45;  Mismatches  114;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |   |     |    | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::| |        ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db         85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
               :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 12
HA1B_BOVIN
ID   HA1B_BOVIN     STANDARD;      PRT;   364 AA.
AC   P13753;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BOLA class I histocompatibility antigen, alpha chain BL3-7 precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258075; PubMed=3133413;
RA   Ennis P.D., Jackson A.P., Parham P.;
RT   "Molecular cloning of bovine class I MHC cDNA.";
RL   J. Immunol. 141:642-651(1988).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to the
CC       immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21043; AAA30641.1; -.
DR   HSSP; P16391; 1ED3.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    364       BOLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN BL3-7.
FT   DOMAIN       28    117       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      118    209       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      210    301       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      302    310       CONNECTING PEPTIDE.
FT   TRANSMEM    311    331
FT   DOMAIN      332    364       CYTOPLASMIC.
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    128    191       BY SIMILARITY.
FT   DISULFID    230    286       BY SIMILARITY.
SQ   SEQUENCE   364 AA;  41513 MW;  622056CF7DCF7873 CRC64;

  Query Match             33.9%;  Score 516;  DB 1;  Length 364;
  Best Local Similarity   38.9%;  Pred. No. 1.2e-35;
  Matches  109;  Conservative   50;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISS 59
              | ||||  | : | |   ||   | |:|||||  || :| :  : | ||| ||:
Db         26 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 84

Qy         60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
              : | : ::  |    :| ||  |:   :| |:  || :| : ||::  |     |:|::|
Db         85 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 144

Qy        118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
              |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db        145 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 203

Qy        178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
              | : ||   |||| ::    |||| || :||: |:: | :: :    :: |  : |
Db        204 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 262

Qy        237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              :||||:| |  | || ||||||||:|:| || :|| : ||
Db        263 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 302


RESULT 13
1A11_HUMAN
ID   1A11_HUMAN     STANDARD;      PRT;   365 AA.
AC   P13746; O19605; O19606; Q29747; Q29835; Q9BCN0; Q9MYI5; Q9TQE9;
AC   Q9TQP6; Q9TQP7;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   HLA class I histocompatibility antigen, A-11 alpha chain precursor
DE   (MHC class I antigen A*11).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*1101 AND A*1102).
RX   MEDLINE=89030641; PubMed=2460344;
RA   Mayer W.E., Jonker M., Klein D., Ivanyi P., van Seventer G.,
RA   Klein J.;
RT   "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT   trans-species mode of evolution.";
RL   EMBO J. 7:2765-2774(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (A*1101 AND A*1102).
RX   MEDLINE=94287401; PubMed=8016845;
RA   Lin L., Tokunaga K., Ishikawa Y., Bannai M., Kashiwase K.,
RA   Kuwata S., Akaza T., Tadokoro K., Shibata Y., Juji T.;
RT   "Sequence analysis of serological HLA-A11 split antigens, A11.1 and
RT   A11.2.";
```

```
RL   Tissue Antigens 43:78-82(1994).
RN   [3]
RP   SEQUENCE OF 26-365 FROM N.A. (A*1101).
RX   MEDLINE=87192928; PubMed=2437024;
RA   Cowan E.P., Jelachich M.L., Biddison W.E., Coligan J.E.;
RT   "DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows
RT   identification of residues involved in epitopes recognized by
RT   antibodies and T cells.";
RL   Immunogenetics 25:241-250(1987).
RN   [4]
RP   SEQUENCE FROM N.A. (A*1103).
RC   TISSUE=Blood;
RX   MEDLINE=20166353; PubMed=10703613;
RA   Tijssen H.J., Sistermans E.A., van den Beucken M.J.G., Krausa P.,
RA   Joosten I.;
RT   "Complete sequence analysis of the A*1103 allele.";
RL   Tissue Antigens 55:68-70(2000).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 2) (A*1103).
RC   TISSUE=Blood;
RX   MEDLINE=20340071; PubMed=10885562;
RA   Tijssen H.J., Sistermans E.A., Joosten I.;
RT   "A unique second donor splice site in the intron 5 sequence of the
RT   HLA-A*11 alleles results in a class I transcript encoding a molecule
RT   with an elongated cytoplasmic domain.";
RL   Tissue Antigens 55:422-428(2000).
RN   [6]
RP   SEQUENCE FROM N.A. (A*1104).
RA   Bettinotti M.P.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 26-206 FROM N.A. (A*1104).
RA   Chandanayingyong D., Sirikong M., Luangtrakool K., Srinak D.,
RA   Rungroung E., Bejchandra S.;
RT   "A11 alleles (A*1104).";
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A. (A*1105).
RX   MEDLINE=99321035; PubMed=10395112;
RA   Morrell G., Whalley J., Stewart A., Day S., Lewis L., Makar Y.,
RA   Ross J., Dunn P.P.;
RT   "Identification of an HLA-A11 serological variant and its
RT   characterization by sequencing based typing.";
RL   Tissue Antigens 53:591-594(1999).
RN   [9]
RP   SEQUENCE OF 26-206 FROM N.A. (A*1105).
RX   MEDLINE=20309230; PubMed=10852390;
RA   Ellis J., Steiner N.K., Kosman C., Henson V., Mitton W., Koester R.,
RA   Ng J., Hartzman R.J., Hurley C.K.;
RT   "Seventeen more novel HLA-A locus alleles.";
RL   Tissue Antigens 55:369-373(2000).
RN   [10]
RP   SEQUENCE FROM N.A. (A*1107).
RX   MEDLINE=21561663; PubMed=11703829;
RA   Pyo C.W., Choi H.B., Han H., Hong Y.S., Kim T.G.;
RT   "Identification of HLA-A*11 variant (A*1107) in the Korean
RT   population.";
RL   Tissue Antigens 58:190-192(2001).
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P13746-1; Sequence=Displayed;
CC       Name=2; Synonyms=Long;
CC         IsoId=P13746-2; Sequence=VSP_008099;
CC         Note=Only produced by allele A*1103;
CC   -!- POLYMORPHISM: The following alleles of A-11 are known: A*1101
CC       (A-11E), A*1102 (A-11K), A*1103, A*1104, A*1105 and A*1107. The
CC       sequence shown is that of A*1101.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X13111; CAA31503.1; -.
DR   EMBL; X13112; CAA31504.1; -.
DR   EMBL; D16841; BAA04117.1; -.
DR   EMBL; D16842; BAA04118.1; -.
DR   EMBL; M16010; AAA65449.1; -.
DR   EMBL; M16007; AAA65449.1; JOINED.
DR   EMBL; M16008; AAA65449.1; JOINED.
DR   EMBL; M16009; AAA65449.1; JOINED.
DR   EMBL; Y17224; CAB38056.1; -.
DR   EMBL; Y17224; CAB38057.1; -.
DR   EMBL; X91399; CAA62745.1; -.
DR   EMBL; U50574; AAB60406.1; -.
DR   EMBL; AF030910; AAB87052.1; -.
DR   EMBL; AF030909; AAB87052.1; JOINED.
DR   EMBL; AF030908; AAB87051.1; -.
DR   EMBL; AF030907; AAB87051.1; JOINED.
DR   EMBL; AJ306733; CAC37336.1; -.
DR   EMBL; AF147455; AAD33991.1; -.
DR   EMBL; AF147454; AAD33991.1; JOINED.
DR   EMBL; AF165065; AAF25781.1; -.
DR   PIR; I83063; I83063.
DR   PIR; S03536; A47636.
DR   HSSP; O19673; 1HSB.
DR   Genew; HGNC:4931; HLA-A.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Signal; Transmembrane; Glycoprotein; Alternative splicing;
KW   Polymorphism.
FT   SIGNAL       1     24
FT   CHAIN       25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-11 ALPHA CHAIN.
FT   DOMAIN      25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN     115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN     207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN     299    308       CONNECTING PEPTIDE.
FT   TRANSMEM   309    332
FT   DOMAIN     333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID   125    188       BY SIMILARITY.
FT   DISULFID   227    283       BY SIMILARITY.
FT   VARSPLIC   337    337       S -> SGGEGVK (in isoform 2).
FT                                /FTId=VSP_008099.
FT   VARIANT     43     43       E -> K (in allele A*1102).
FT                                /FTId=VAR_004353.
FT   VARIANT    133    133       F -> L (in allele A*1107).
FT                                /FTId=VAR_016731.
FT   VARIANT    168    168       K -> E (in allele A*1105).
FT                                /FTId=VAR_016732.
FT   VARIANT    175    175       H -> R (in allele A*1103).
FT                                /FTId=VAR_016733.
FT   VARIANT    176    176       A -> E (in allele A*1103).
```

```
FT                                /FTId=VAR_016734.
FT   VARIANT     187    187       R -> T (in allele A*1104).
FT                                /FTId=VAR_016735.
FT   VARIANT     345    345       T -> S (in allele A*1105).
FT                                /FTId=VAR_016736.
SQ   SEQUENCE   365 AA;  40937 MW;  FE449CE2D4BF6CC5 CRC64;

  Query Match             33.8%;  Score 514;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 1.8e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | :   |    |   | |:|||||  || :| :  |:|:|||  ||:      : |
Db         26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|       ||  |:   :| |:: |||:|:: ||::  |      || :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
             :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    ||||| || :|| ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |||||||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
1A03_HUMAN
ID   1A03_HUMAN     STANDARD;      PRT;   365 AA.
AC   P04439; O19546; O19756; Q9GJE6; Q9GJE7; Q9GJE8; Q9MYG4; Q9TPR8;
DT   13-AUG-1987 (Rel. 05, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   HLA class I histocompatibility antigen, A-3 alpha chain precursor
DE   (MHC class I antigen A*3).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*0301).
RX   MEDLINE=84207948; PubMed=6609814;
RA   Strachan T., Sodoyer R., Damotte M., Jordan B.R.;
RT   "Complete nucleotide sequence of a functional class I HLA gene,
RT   HLA-A3: implications for the evolution of HLA genes.";
RL   EMBO J. 3:887-894(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (A*0301).
RA   Ellexson M.E., Hildebrand W.H.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (A*0301).
RA   Shiina S., Tamiya G., Oka A., Inoko H.;
RT   "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (A*0302).
RX   MEDLINE=85290871; PubMed=2993417;
RA   Cowan E.P., Jordan B.E., Coligan J.E.;
RT   "Molecular cloning and DNA sequence analysis of genes encoding
RT   cytotoxic T lymphocyte-defined HLA-A3 subtypes: the E1 subtype.";
RL   J. Immunol. 135:2835-2841(1985).
RN   [5]
RP   SEQUENCE FROM N.A. (A*0302).
RA   Bettinotti M.P., Hadzikadic L., Adams S., Marincola F.M.;
RT   "Complete coding sequence of HLA-A*0302.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1-341 FROM N.A. (A*0304).
RX   MEDLINE=99180630; PubMed=10079303;
RA   Santos S., Balas A., Garcia-Sanchez F., Lillo R., Merino J.L.,
RA   Vicario J.L.;
RT   "Complete cDNA coding sequence of a new HLA-A3 subtype (A*0304) with a
RT   new HLA polymorphism at exon 3.";
RL   Immunogenetics 49:360-361(1999).
RN   [7]
RP   SEQUENCE OF 26-206 FROM N.A. (A*0305).
RA   Becher M.P., Wu J., Williams T.;
RT   "Novel human HLA-A alleles identified in potential bone marrow
RT   donors.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 26-298 FROM N.A. (A*0305).
RC   TISSUE=Blood;
RA   Poli F., Frison S., Crespiatico L., Longhi E.;
RT   "Identification of a HLA-A*03 new variant.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to
CC       the immune system.
CC   -!- SUBUNIT: Dimer of alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- POLYMORPHISM: The following alleles of A-3 are known: A*0301
CC       (A-3.1), A*0302, A*0304 and A*0305. The sequence shown is that of
CC       A*0301.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00492; CAA25162.1; ALT_SEQ.
DR   EMBL; U32184; AAB63980.1; -.
DR   EMBL; AP000520; BAB63400.1; -.
DR   EMBL; AP000519; BAB63400.1; JOINED.
DR   EMBL; AF217561; AAF28734.1; -.
DR   EMBL; AF015930; AAB66582.1; -.
DR   EMBL; AF190719; AAF03243.1; -.
DR   EMBL; AF190718; AAF03243.1; JOINED.
DR   EMBL; AJ401085; CAC06086.1; -.
DR   EMBL; AJ401086; CAC06087.1; -.
DR   EMBL; AJ401087; CAC06088.1; -.
DR   PIR; A02192; HLHUA3.
DR   HSSP; O19673; 1HSB.
DR   Genew; HGNC:4931; HLA-A.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-3 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
```

```
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
FT   VARIANT      176    176       E -> V (in allele A*0302).
FT                                 /FTId=VAR_004351.
FT   VARIANT      180    180       L -> Q (in allele A*0302).
FT                                 /FTId=VAR_004352.
FT   VARIANT      185    185       D -> E (in allele A*0305).
FT                                 /FTId=VAR_016604.
FT   VARIANT      199    199       G -> R (in allele A*0304).
FT                                 /FTId=VAR_016605.
FT   CONFLICT     319    319       G -> A (IN REF. 6).
SQ   SEQUENCE   365 AA;  40840 MW;  DEDFCEC4450E0580 CRC64;

  Query Match             33.7%;  Score 512;  DB 1;  Length 365;
  Best Local Similarity   39.6%;  Pred. No. 2.6e-35;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |     |    | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               | ::::|        ||   |:   :| |:  |||:|:: ||:: |     ||  :  |||
Db         85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 202

Qy        180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db        203 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 261

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||:| |  : || |||||||| |:| || :|| : ||
Db        262 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
1A80_HUMAN
ID   1A80_HUMAN     STANDARD;      PRT;   365 AA.
AC   Q09160;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   HLA class I histocompatibility antigen, A-80 alpha chain precursor
DE   (MHC class I antigen A*80) (Aw-80) (A-1).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*8001).
RX   MEDLINE=94245293; PubMed=8188325;
RA   Balas A., Garcia-Sanchez F., Gomez-Reino F., Vicario J.L.;
RT   "Characterization of a new and highly distinguishable HLA-A allele in
RT   a Spanish family.";
RL   Immunogenetics 39:452-452(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (A*8001).
RA   Domena J.D.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Involved in the presentation of foreign antigens to the
CC       immune system.
CC   -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-
CC       microglobulin).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- POLYMORPHISM: The only allele of A-80 known is A*8001 which is
CC       shown here.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U03754; AAC04322.1; -.
DR   EMBL; L18898; AAA17012.1; -.
DR   PIR; I59638; I38439.
DR   HSSP; Q95352; 1HHK.
DR   Genew; HGNC:4931; HLA-A.
DR   MIM; 142800; -.
DR   GO; GO:0005887; C:integral to plasma membrane; NAS.
DR   GO; GO:0030106; F:MHC class I receptor activity; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     24
FT   CHAIN       25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-80 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   365 AA;  40791 MW;  CE1BC1CD60CA8FA8 CRC64;

  Query Match             33.6%;  Score 510;  DB 1;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 3.8e-35;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMW 62
              |||: | |    |     |    | |:|||||  || :| :  |:|:||| ||:      : |
Db         26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy         63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
               : ::::|        :  |:   :| |:: |||:|:: ||:: |     ||  :  |||
Db         85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              :|::    |   | ||:  |  || :||  : || | |||||  :|   |:: || |:  |
Db        145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy        181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
              :  ||   :||| ::    |||| ||::||  ||: | :| :    :: |  :  | ||
Db        204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy        240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||:| |  : || |:|:|||| |:| || :|| : ||
Db        263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


Search completed: May  4, 2004, 11:35:40
Job time : 8.33333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:32:18 ; Search time 34.6667 Seconds
(without alignments)
2512.010 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1220	80.3	268	4	Q86WL1	Q86wl1 homo sapien
2	1140	75.0	358	11	Q8C2A6	Q8c2a6 mus musculu
3	1140	75.0	359	11	Q9D754	Q9d754 mus musculu
4	802	52.8	272	11	Q9R105	Q9r105 rattus norv
5	592	38.9	116	4	Q9HC69	Q9hc69 homo sapien
6	547.5	36.0	359	7	Q8HX81	Q8hx81 ornithorhyn
7	543.5	35.8	340	7	Q9BD50	Q9bd50 pongo pygma
8	542.5	35.7	334	7	Q9TQK3	Q9tqk3 homo sapien
9	542.5	35.7	341	4	Q9NPL2	Q9npl2 homo sapien
10	542.5	35.7	341	7	Q95460	Q95460 homo sapien
11	542.5	35.7	341	7	Q9BCU3	Q9bcu3 pan troglod
12	540.5	35.6	354	7	Q95HB3	Q95hb3 anas platyr
13	539.5	35.5	341	7	Q9BCU4	Q9bcu4 pan troglod
14	530	34.9	105	4	Q9HC71	Q9hc71 homo sapien
15	521	34.3	356	7	Q8HX66	Q8hx66 sus scrofa
16	520	34.2	332	7	Q30990	Q30990 pan troglod
17	520	34.2	365	7	Q9TPL7	Q9tpl7 pan troglod
18	519	34.1	312	7	Q860B4	Q860b4 homo sapien
19	517	34.0	352	7	Q8MHT1	Q8mht1 sus scrofa
20	517	34.0	364	7	O19243	O19243 sus scrofa
21	514	33.8	273	7	Q95IG6	Q95ig6 homo sapien
22	514	33.8	352	7	Q8SPA9	Q8spa9 sus scrofa
23	514	33.8	361	7	Q8HX63	Q8hx63 sus scrofa
24	514	33.8	364	7	Q8HX61	Q8hx61 sus scrofa
25	513	33.8	330	7	O19356	O19356 macaca mula
26	513	33.8	331	7	O02944	O02944 macaca mula
27	513	33.8	333	7	O98030	O98030 papio anubi
28	513	33.8	333	7	O98031	O98031 papio anubi
29	512	33.7	129	4	Q9UK37	Q9uk37 homo sapien
30	512	33.7	330	7	O02947	O02947 macaca mula
31	512	33.7	330	7	O02946	O02946 macaca mula
32	511	33.6	331	7	O02945	O02945 macaca mula
33	511	33.6	357	7	Q30886	Q30886 pan paniscu
34	511	33.6	363	7	Q9MXI5	Q9mxi5 pan troglod
35	511	33.6	363	7	Q9MWK4	Q9mwk4 gorilla gor
36	511	33.6	365	7	Q9MXI6	Q9mxi6 pan troglod
37	511	33.6	365	7	Q9MXM7	Q9mxm7 pan troglod
38	510	33.6	360	7	Q9GJ24	Q9gj24 homo sapien
39	510	33.6	364	7	Q8SPA4	Q8spa4 sus scrofa
40	510	33.6	365	7	Q30900	Q30900 pan paniscu
41	510	33.6	365	7	Q8HWQ9	Q8hwq9 homo sapien
42	509.5	33.5	360	7	Q95558	Q95558 peromyscus
43	509	33.5	273	7	Q9TQP8	Q9tqp8 homo sapien
44	509	33.5	298	7	Q8MHN8	Q8mhn8 homo sapien
45	509	33.5	318	7	Q7YPW4	Q7ypw4 homo sapien

ALIGNMENTS

RESULT 1
Q86WL1
ID Q86WL1 PRELIMINARY; PRT; 268 AA.
AC Q86WL1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemochromatosis (Fragment).
GN HFE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kutlar F., Nechtman J., Leithner C.;
RT "Direct isolation of hemochromatosis (HFE) mRNA from the whole blood
RT of a normal Caucasian individual.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY205604; AAO47091.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1 1
SQ SEQUENCE 268 AA; 30952 MW; D725DE42AC08DAA5 CRC64;

Query Match 80.3%; Score 1220; DB 4; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.4e-106;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         59 SQMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGY 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SQMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGY 60

Qy        119 DGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRG 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRG 120

Qy        179 VLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 180

Qy        239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||||
Db        181 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 218


RESULT 2
Q8C2A6
ID   Q8C2A6      PRELIMINARY;      PRT;   358 AA.
AC   Q8C2A6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemochromatosis.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK088986; BAC40688.1; -.
DR   PIR; PT0706; PT0706.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
SQ   SEQUENCE   358 AA;  40421 MW;  EE88FB6E5AAC844D CRC64;

  Query Match             75.0%;  Score 1140;  DB 11;  Length 358;
  Best Local Similarity   72.2%;  Pred. No. 6.7e-99;
  Matches  203;  Conservative   30;  Mismatches   40;  Indels    8;  Gaps    1;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              ||||| |||||||| |||| |||| ||||||||| |: |||| ||| ||:  : |||:||
Db         29 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 88

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
              |||||||||:|| |||||||  |:|||         ||| |||:||||: ||||| |:|:
Db         89 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 148

Qy        116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
              |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db        149 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 208

Qy        176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
              ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
```

```
Db        209 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 268

Qy        236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||| ||||:|||| ||:| |:|||||||||||||   ||
Db        269 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 309


RESULT 3
Q9D754
ID   Q9D754      PRELIMINARY;      PRT;   359 AA.
AC   Q9D754;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Adult male tongue cDNA, RIKEN full-length enriched library,
DE   clone:2310032M04, full insert sequence.
GN   HFE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Tongue;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AK009581; BAB26373.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   MGD; MGI:109191; Hfe.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   359 AA;  40534 MW;  586657B7F9FF20B4 CRC64;

  Query Match             75.0%;  Score 1140;  DB 11;  Length 359;
  Best Local Similarity   72.2%;  Pred. No. 6.8e-99;
  Matches  203;  Conservative   30;  Mismatches   40;  Indels    8;  Gaps    1;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
```

```
                |||||  |||||||| |||| |||| |||||||| |: |||| ||| ||:  : |||:||
Db          30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy          64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
               |||||||||:|| ||||||| |:|||         ||| |||:|||: ||||| |:|:
Db          90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy         116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
               |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db         150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy         176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
               |||||  |||| |||||  |   |: |:|||:||:::||||||:|||| ||:|||:  |: ||
Db         210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy         236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
               ||||  |||||:|||| ||:| |:|||||||||||||    ||
Db         270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310


RESULT 4
Q9R105
ID   Q9R105      PRELIMINARY;      PRT;   272 AA.
AC   Q9R105;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hemochromatosis gene product HFE splice variant delE2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Testis;
RA   Liew Y.-F., Shaw N.-S.;
RT   "Alternative splice variant of the hemochromatosis gene HFE in iron
RT   overloaded rats.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF176534; AAD49965.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   272 AA;  30757 MW;  1D91063CCBEF5502 CRC64;

  Query Match             52.8%;  Score 802;  DB 11;  Length 272;
  Best Local Similarity   75.1%;  Pred. No. 3.1e-67;
  Matches  139;  Conservative   22;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
               ||| |||||||:  ||||| |:|||||||||||||| ||:| ||||||| ||:||| |:
Db          39 ESHILQVILGCEVHEDNSTSGFWKYGYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHR 98

Qy         152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQ 211
               |||||:| ||:|||| ||:|:||| |||| |||| |||||  |  |: |:|||:|||::||
Db          99 IRARQSRDYLQRDCPQQLKQVLELQRGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQ 158


Qy         212 NITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPL 271
               ||||:|||| ||:|||:  |::|||||||||||:|||| |||| |::|||||||||||
Db         159 NITMRWLKDSQPLDAKDVNPENVLPNGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPL 218

Qy         272 IVIWE 276
                 ||
Db         219 TATWE 223


RESULT 5
Q9HC69
ID   Q9HC69      PRELIMINARY;      PRT;   116 AA.
AC   Q9HC69;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hemochromatosis splice variant 861-2305del (Fragment).
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF144241; AAG29576.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   116 AA;  13541 MW;  AC0333B096A3F47B CRC64;

  Query Match             38.9%;  Score 592;  DB 4;  Length 116;
  Best Local Similarity   99.1%;  Pred. No. 6.2e-48;
  Matches  106;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          89 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 60

Qy         149 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVT 195
               |||||||||||||||||||||||||||||||||||||||||||||:
Db          61 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVS 107


RESULT 6
Q8HX81
ID   Q8HX81      PRELIMINARY;      PRT;   359 AA.
AC   Q8HX81;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MHC class I antigen.
OS   Ornithorhynchus anatinus (Duckbill platypus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX   NCBI_TaxID=9258;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22226695; PubMed=12242589;
RA   Miska K.B., Harrison G.A., Hellman L., Miller R.D.;
RT   "The major histocompatibility complex in monotremes: an analysis of
RT   the evolution of Mhc class I genes across all three mammalian
RT   subclasses.";
RL   Immunogenetics 54:381-393(2002).
```

```
DR   EMBL; AY112715; AAM54213.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
SQ   SEQUENCE   359 AA;  39847 MW;  3FC701E44FD53B66 CRC64;

  Query Match             36.0%;  Score 547.5;  DB 7;  Length 359;
  Best Local Similarity   41.7%;  Pred. No. 3.9e-43;
  Matches  116;  Conservative   49;  Mismatches  106;  Indels    7;  Gaps    6;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFV-FYDDESR-RVEPRTPWVSSRISSQM 61
              |||||  |  :  |  |     |:|  |  |:||||:  ||  |   |  |  |  ||||||:      :
Db         25 RSHSLRYFYTGVSRPGSGVSEFTAVGYVDEVQFVRFEGDGGRGRAEPRTPWIRDNEGQEY 84

Qy         62 WLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNST-EGYWKYGYD 119
              |  :  :|  ||     :|   :  |  :  :| |||  ||| |  :||:::||::   :|: :||||
Db         85 WDRNTQILKDTAQVFRTNLQTALGYYNQSKEGSHTFQRMYGCQLREDDTPGQGFTQYGYD 144

Qy        120 GQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
              |:|::      ||  |  ||:  ||  :||  :   | :::||||  :|    |:: :  |
Db        145 GRDYIALDKATLSWTAADAGALNTKRKWEADRTIAERDKAYLEDECIEWLRKYVGFGGPS 204

Qy        180 LDQQVPPLVKVTHHV--TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPN 237
              |  :  ||  |:|| |         :||||||  :|| :|  ||  ::  :  | ::|  |      |:
Db        205 LTRADPPEVEVTRHTGPDGEDVSLRCRALGFYPADIKFKWEREGKDM-SQEMEFVGTRPS 263

Qy        238 GDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              |||  :|  |  ::  ||  ||||:|  |  |||  ||  |||  |  |
Db        264 GDGNFQKWASVNVPRGEEQKYVCVVEHEGLGQPLAVKW 301


RESULT 7
Q9BD50
ID   Q9BD50      PRELIMINARY;      PRT;   340 AA.
AC   Q9BD50;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MHC class I related protein MR1 isoform.
GN   MR1.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.,
RA   Martinez-Naves E.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ271828; CAC28215.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   340 AA;  39375 MW;  A893952B78725F17 CRC64;

  Query Match             35.8%;  Score 543.5;  DB 7;  Length 340;
  Best Local Similarity   39.5%;  Pred. No. 8.7e-43;
  Matches  107;  Conservative   51;  Mismatches  110;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:|||  |  :|  |:   |:  |  ::||||      ||  ::: |||  ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIRGVPEFISVGYVDSHPITTYDSVTQQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMPTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :|  |:||   || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  |||  | |  :   |   |  || ::   :  : :|| :|  | |::  || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKRAWEANQHELQYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|     |   ||||  |:|  :||  |  | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTTLFCKAHGFYPPEIYMTWMKNGEEI-VQEMDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | |  :  : |     |:| |||  |:   | |
Db        261 QTWASFELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 8
Q9TQK3
ID   Q9TQK3      PRELIMINARY;      PRT;   334 AA.
AC   Q9TQK3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MHC class I-related protein MR1 (Fragment).
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=99003494; PubMed=9784382;
RA   Yamaguchi H., Kurosawa Y., Hashimoto K.;
RT   "Expanded genomic organization of conserved mammalian MHC class I-
RT   related genes, human MR1 and its murine ortholog.";
RL   Biochem. Biophys. Res. Commun. 250:558-564(1998).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF073485; AAC72900.1; -.
DR   EMBL; AF073484; AAC72900.1; JOINED.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
```

```
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   334 AA;  38586 MW;  4C3E3A8248A39BA4 CRC64;

  Query Match             35.7%;  Score 542.5;  DB 7;  Length 334;
  Best Local Similarity   39.5%;  Pred. No. 1.1e-42;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||| |      ||  :|: ||| ||::  ::   |
Db         16 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 75

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::||| ||| | ::|||: || || |: :| |||||
Db         76 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 134

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |  || ::      : :|| :| | |:: || |:  | :
Db        135 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 194

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        195 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 253

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        254 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 284


RESULT 9
Q9NPL2
ID   Q9NPL2      PRELIMINARY;      PRT;   341 AA.
AC   Q9NPL2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mr1 protein.
GN   MR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peripheral blood;
RX   MEDLINE=20470599; PubMed=11019920;
RA   Parra-Cuadrado J.F., Navarro P. ., Mirones I., Setien F., Oteo M.,
RA   Martinez-Naves E.;
RT   "A study on the polymorphism of human MHC class I-related MR1 gene and
RT   identification of an MR1-like pseudogene.";
RL   Tissue Antigens 56:170-172(2000).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ249778; CAB77667.1; -.
DR   PIR; A57136; A57136.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match             35.7%;  Score 542.5;  DB 4;  Length 341;
  Best Local Similarity   39.5%;  Pred. No. 1.1e-42;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::|||| |      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::||| ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |  || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 10
Q95460
ID   Q95460      PRELIMINARY;      PRT;   341 AA.
AC   Q95460;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Class I histocompatibility antigen-like protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RX   MEDLINE=95350662; PubMed=7624800;
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "A gene outside the human MHC related to classical HLA class I
RT   genes.";
RL   Science 269:693-695(1995).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; U22963; AAC50174.1; -.
DR   PIR; A57136; A57136.
DR   HSSP; Q30201; 1A6Z.
DR   Genew; HGNC:4975; MR1.
DR   GO; GO:0030106; F:MHC class I receptor activity; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   341 AA;  39366 MW;  2990C1F3F0A1CAD9 CRC64;

  Query Match             35.7%;  Score 542.5;  DB 7;  Length 341;
  Best Local Similarity   39.5%;  Pred. No. 1.1e-42;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;
```

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:|||  |  :|  |:    |:  | ::||||      ||  :|: |||  ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :|  |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |  |  ||| | |  :  |   |  || ::      : :|| :| | |:: || |: :| :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291
```

RESULT 11
Q9BCU3
ID Q9BCU3 PRELIMINARY; PRT; 341 AA.
AC Q9BCU3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I related protein, MR1B1 isoform.
GN MR1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Naves E.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT "Identification of MR1 cDNA sequences in non-human primates.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AJ275984; CAC34272.1; -.
DR HSSP; Q30201; 1A6Z.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
FT VARIANT 197 197 I -> T (IN REF. 2).
SQ SEQUENCE 341 AA; 39394 MW; FBFA822BCAB2C7A8 CRC64;

Query Match 35.7%; Score 542.5; DB 7; Length 341;
Best Local Similarity 39.5%; Pred. No. 1.1e-42;
Matches 107; Conservative 51; Mismatches 110; Indels 3; Gaps 3;

```
Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:|||  |  :|  |:    |:  | ::||||      ||  :|: |||  ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :|  |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              |  |  ||| | |  :  |   |  || ::      : :|| :| | |:: || |: :| :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
               ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| ||| |:   | |
Db        261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291
```

RESULT 12
Q95HB3
ID Q95HB3 PRELIMINARY; PRT; 354 AA.
AC Q95HB3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen alpha chain (Fragment).
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF393511; AAK84356.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC.
FT NON_TER 1 1
SQ SEQUENCE 354 AA; 39405 MW; D44FACEE7B60468F CRC64;

Query Match 35.6%; Score 540.5; DB 7; Length 354;
Best Local Similarity 37.3%; Pred. No. 1.8e-42;
Matches 101; Conservative 60; Mismatches 109; Indels 1; Gaps 1;

```
Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWLQL 65
              |||  |     |:   |:  |  |:|||| ::|| || |:|  :|||  |::  :  | |
Db         24 HSLRYFATAVSDPSPGVPQFVAVGYVDGEVFVRYDSETRGMEPRVDWIADNMDQQYWNGE 83

Qy         66 SQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLE 125
              :::|:|  : :: ||   |: | :|  |: ||||| : ||:: || |  |:::|||:|:: :
Db         84 TENLRGAEQIYRVDLETLRERYNQSRGSHTLQHMFGCDLLEDRSISGFFQYGYEGREFIA 143

Qy        126 FCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVP 185
                ||  : ||:  |  || :||      | : : |||  |    |:: :  |: ||:::
Db        144 LDKDTWTFTAADAAAQITKRKWEEDGTVAERRKYYLENTCIEWLRKYVRYGKDVLERRER 203

Qy        186 PLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGW 245
              | |:|:      : :| |||  :||: |:: ||||   :  :| :   :|| |||| |
Db        204 PEVRVSGMEADKILSLSCRAHGFYPRPISISWLKDGM-VQEQETQRGSTVPNSDGTYHIW 262
```

```
Qy        246 ITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |: | ||:: :| |:|||  | || :  ||
Db        263 ATIDVVPGDKDKYQCRVEHASLPQPGLFSWE 293


RESULT 13
Q9BCU4
ID   Q9BCU4      PRELIMINARY;      PRT;   341 AA.
AC   Q9BCU4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MHC class I related protein, MR1B1 isoform.
GN   MR1.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martinez-Naves E.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ275982; CAC34274.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
FT   VARIANT     197    197       T -> I (IN REF. 2).
SQ   SEQUENCE   341 AA;  39382 MW;  DFF16AF1FAB2D272 CRC64;

  Query Match           35.5%;  Score 539.5;  DB 7;  Length 341;
  Best Local Similarity  39.5%;  Pred. No. 2.1e-42;
  Matches  107;  Conservative   50;  Mismatches  111;  Indels    3;  Gaps    3;

Qy          4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQMWL 63
              |:||| |  :| |:   |:  | ::||||      ||  :|: ||| ||::  ::   |
Db         23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
              : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db         83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy        124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
              | |  ||| | | :  |   | || ::      : :|| :| | |:: || |:  | :
Db        142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy        184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
              ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db        202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy        243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
              | | :: : |     |:| |||  |:   | |

Db        261 QTWASVELDPQSSNLYSCHVEHCGVHMVLQV 291


RESULT 14
Q9HC71
ID   Q9HC71      PRELIMINARY;      PRT;   105 AA.
AC   Q9HC71;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hemochromatosis splice variant 838-2283del (Fragment).
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF144239; AAG29574.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   105 AA;  12233 MW;  4A50E52AA275D4B0 CRC64;

  Query Match           34.9%;  Score 530;  DB 4;  Length 105;
  Best Local Similarity  94.0%;  Pred. No. 3.7e-42;
  Matches   94;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         89 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 148
              |:||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db          1 HTKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRASEPRAWPTKLEWE 60

Qy        149 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLV 188
              ||:|||| ||||||||||||||||||||||||||||| |
Db         61 GHKVRARQNGAYLERDCPAQLQQLLELGRGVLDQQVPPKV 100


RESULT 15
Q8HX66
ID   Q8HX66      PRELIMINARY;      PRT;   356 AA.
AC   Q8HX66;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MHC class I antigen (Fragment).
GN   SLA-1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Martens G.W., Baker J.E., Smith D.M.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY135589; AAN35107.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0006955; P:immune response; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
```

```
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   356 AA;  39585 MW;  94FC7A461DBF555B CRC64;

  Query Match             34.3%;  Score 521;  DB 7;  Length 356;
  Best Local Similarity   39.9%;  Pred. No. 1.2e-40;
  Matches  110;  Conservative   48;  Mismatches  110;  Indels    8;  Gaps    7;

Qy          6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDE--SRRVEPRTPWVSSRISSQMWL 63
              ||| | :   |  ||| | | |:|||||  || :| :  : |:||| ||:      : |
Db         19 HSLRYFYTAVSRPDLGDSRFIAVGYVDDTQFVRFDSDAPNPRMEPRAPWIQQE-GQEYWD 77

Qy         64 QLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDN-STEGYWKYGYDGQ 121
              : :::  |      ||  |:   :| |:  |||:| : ||::  |      || :  |||
Db         78 EETRNAMGSAQNDRVDLKTLRGYYNQSEAGSHTIQRMYGCDVGPDGLLLRGYDQDAYDGA 137

Qy        122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
              |::    |   | ||:  |  || :||   : | | |:|||  |   ||: ||:|:  |
Db        138 DYIALNEDLRSWTAADTAAQITKRKWEAANV-AEQERSYLEGTCVEWLQKYLEMGKDTLQ 196

Qy        182 QQVPPLVKVTHHVTSSV-TTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              :  ||   || | :| :  |||| || :||: |:: | :: |   ::: |  :  |:|||
Db        197 RAEPPKTHVTRHPSSDLGVTLRCWALGFYPKEISLTWQREGQD-QSQDMELVETRPSGDG 255

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |:| |  | |||||| ||| |:| || :|| : |:
Db        256 TFQKWAALVVPPGEEQSYTCHVQHEGLQEPLTLRWD 291

Search completed: May  4, 2004, 11:38:32
Job time : 34.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:32:18 ; Search time 49.3333 Seconds
(without alignments)
1580.739 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1520	100.0	276	2	AAW94296	Aaw94296 HFE mutan
2	1520	100.0	276	6	ABG72686	Abg72686 Human hae
3	1520	100.0	276	6	ABU62093	Abu62093 HFE mutan
4	1520	100.0	348	4	AAB36871	Aab36871 Human her
5	1513	99.5	276	2	AAW94295	Aaw94295 Wild-type
6	1513	99.5	276	6	ABG72685	Abg72685 Human hae
7	1513	99.5	276	6	ABU62092	Abu62092 HFE mutan
8	1513	99.5	348	2	AAW36499	Aaw36499 Hereditar
9	1513	99.5	348	3	AAB19149	Aab19149 A human h
10	1513	99.5	348	4	AAB36869	Aab36869 Human her
11	1509	99.3	348	4	AAB36872	Aab36872 Human her
12	1508	99.2	438	5	AAU80035	Aau80035 Beta 2 mi
13	1506	99.1	276	6	ABU62091	Abu62091 HFE polyp
14	1502	98.8	348	4	AAB36870	Aab36870 Human her
15	1493	98.2	276	2	AAW94297	Aaw94297 HFE mutan
16	1493	98.2	276	6	ABG72687	Abg72687 Human hae
17	523	34.4	361	4	AAB36873	Aab36873 Rabbit le
18	514	33.8	92	6	ABP68379	Abp68379 Human col
19	514	33.8	365	4	AAB36874	Aab36874 MHC class
20	506	33.3	274	3	AAY68275	Aay68275 Human leu
21	506	33.3	274	3	AAY52929	Aay52929 HLA-A2/A2
22	506	33.3	274	4	AAB58690	Aab58690 HLA-A2/A2
23	506	33.3	280	4	AAU10225	Aau10225 Human leu
24	506	33.3	280	6	ABU08672	Abu08672 Human his
25	506	33.3	415	4	AAU10224	Aau10224 Human par
26	506	33.3	415	6	ABU08671	Abu08671 Human sin
27	506	33.3	510	6	AAE36053	Aae36053 B2M-aTacV
28	505	33.2	365	3	AAY68265	Aay68265 Human leu
29	505	33.2	365	3	AAY52919	Aay52919 HLA-A2/A2
30	505	33.2	365	4	AAB58680	Aab58680 HLA-A2/A2
31	505	33.2	368	4	AAM24017	Aam24017 Human EST
32	504	33.2	274	3	AAY68276	Aay68276 Human leu
33	504	33.2	274	3	AAY52930	Aay52930 HLA-A2/A2
34	504	33.2	274	4	AAB58691	Aab58691 HLA-A2/A2
35	504	33.2	365	3	AAY68268	Aay68268 Human leu
36	504	33.2	365	3	AAY52922	Aay52922 HLA-A2/A2
37	504	33.2	365	4	AAB58683	Aab58683 HLA-A2/A2
38	503	33.1	274	1	AAP80911	Aap80911 Consensus
39	503	33.1	365	3	AAY68267	Aay68267 Human leu
40	503	33.1	365	3	AAY52921	Aay52921 HLA-A2/A2
41	503	33.1	365	4	AAB58682	Aab58682 HLA-A2/A2
42	502	33.0	274	3	AAY68274	Aay68274 Human leu
43	502	33.0	274	3	AAY52928	Aay52928 HLA-A2/A2
44	502	33.0	274	4	AAB58689	Aab58689 HLA-A2/A2
45	502	33.0	365	3	AAY68266	Aay68266 Human leu

ALIGNMENTS

RESULT 1
AAW94296
ID AAW94296 standard; peptide; 276 AA.
XX
AC AAW94296;
XX
DT 27-APR-1999 (first entry)
XX
DE HFE mutant (H63D-HFE) polypeptide sequence.
XX
KW HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW transfusion; protein replacement therapy; hereditary hemochromatosis;
KW transferrin receptor; iron deficiency; anemia; mutant.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "indicated in the sequence listing as Arg"
FT Misc-difference 41
FT /label= H63D
FT /note= "wild type His (of the mature protein sequence) is
FT replaced by Asp"
XX
PN WO9856814-A1.
XX
PD 17-DEC-1998.
XX
PF 12-JUN-1998; 98WO-US012436.
XX
PR 13-JUN-1997; 97US-00876010.
XX
PA (PROG-) PROGENITOR INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 1999-080886/07.
XX
PT New treatment of an iron overload disease - comprises use of HFE
PT polypeptides provided in a complex with full length, wild type human
PT (2m), useful in protein replacement therapy.
XX
PS Claim 3; Page 14; 36pp; English.
XX
CC The present sequence represents a H63D-HFE mutant polypeptide. The HFE
CC polypeptides (AAW94295-297) provided in a complex with full length, wild
CC type human beta-2-microglobulin (beta2m) form compositions in the

CC treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC other iron overload conditions resulting from secondary causes (e.g.
CC repeated transfusions). Data regarding the structure and function
CC correlations of HFE polypeptides is useful in designing drugs that
CC modulate the HFE gene and HFE activity. The polypeptides are also useful
CC in protein replacement therapy for individuals possessing a defective HFE
CC gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the polypeptides
CC are also useful in treating primary and secondary iron overload diseases.
CC The modulators of the transferrin receptor are useful in treating iron
CC deficiency conditions such as anemia, and in modulating the amount of
CC iron transported into a cell. The HFE polypeptides provide a molecular
CC basis for the relationship between HFE and iron metabolism, which enables
CC treatment of iron overload and deficiency diseases
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1520; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.6e-133;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

RESULT 2
ABG72686
ID ABG72686 standard; protein; 276 AA.
XX
AC ABG72686;
XX
DT 05-MAR-2003 (first entry)
XX
DE Human haemochromatosis (HFE) mature protein, mutant H41D.
XX
KW Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW Beta2m; transferrin receptor; anaemia; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 41
FT /note= "Wild-type His substituted by Asp"
XX
PN US6391852-B1.
XX
PD 21-MAY-2002.
XX
PF 12-JUN-1998; 98US-00094964.
XX
PR 13-JUN-1997; 97US-00876010.
XX
PA (BIRA) BIO-RAD LAB INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-155377/15.
XX
PT Method of treating an iron overload disease comprises administration of a
PT soluble complex comprising a 276 amino acid HFE polypeptide and a full
PT length, wild-type human beta2m.
XX
PS Claim 2; Col 2; 17pp; English.
XX
CC The invention relates to a method of treating an iron overload disease
CC comprising administration of a soluble complex comprising a 276 amino
CC acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC (ABG72685-ABG72687) and a full length, wild-type human beta2m (beta2-
CC microglobulin). In a HeLa cell based assay, binding and uptake of ^1^2^5I
CC -transferrin in the presence of purified H63D-HFE/beta2m heterodimers was
CC determined. At a concentration of 250 nM H63D-HFE/ beta2m heterodimers,
CC the transferrin receptor (TfR) displayed a KD for transferrin of 28 nM.
CC At the same concentration of normal HFE/beta_2m heterodimers, TfR
CC displayed a KD for transferrin of 40 nM. In the absence of any
CC HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of 7nM. It
CC was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC efficient in decreasing TfR affinity for transferrin compared to wild-
CC type HFE. The method is useful for treating iron overload diseases and
CC iron deficiency e.g. anaemia. The present sequence is the H63D (residue
CC 63 of the full length protein, 41 of the mature form) mutant from of
CC mature HFE used to investigate the role of the His residue in transferrin
CC receptor binding to transferrin
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1520; DB 6; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.6e-133;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

RESULT 3
ABU62093
ID ABU62093 standard; protein; 276 AA.
XX
AC ABU62093;
XX
DT 01-OCT-2003 (first entry)
XX
DE HFE mutant polypeptide #2 useful for treating iron diseases.
XX
KW Iron overload disease; iron deficiency disease; HFE polypeptide;
KW beta2 microglobulin; beta2m; hereditary haemochromatosis; HH; anaemia;
KW protein replacement therapy; defective HFE gene; human; antianaemic;
KW mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.

XX
PN US2003073627-A1.
XX
PD 17-APR-2003.
XX
PF 04-MAR-2002; 2002US-00092404.
XX
PR 13-JUN-1997; 97US-00876010.
PR 12-JUN-1998; 98US-00094964.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-567313/53.
XX
PT Treating an iron overload disease (e.g. hemochromatosis) or an iron
PT deficiency disease (e.g. anemia), comprises administering to a patient an
PT HFE polypeptide and full-length, wild type human beta-2 microglobulin.
XX
PS Claim 5; Page 1; 14pp; English.
XX
CC The present invention relates to a method for treating iron overload
CC diseases and iron deficiency diseases. The method comprises administering
CC to a patient an HFE polypeptide. The HFE polypeptide is provided in a
CC complex with full-length, wild type human beta2 microglobulin (beta2m).
CC The method and HFE polypeptide are useful for diagnosing or treating an
CC iron overload disease (e.g. hereditary haemochromatosis, HH) or an iron
CC deficiency disease (e.g. anaemia). The HFE polypeptide is also useful in
CC protein replacement therapy for individuals having a defective HFE gene.
CC The present sequence represents an HFE polypeptide useful for treating
CC iron diseases
XX
SQ Sequence 276 AA;

Query Match 100.0%; Score 1520; DB 6; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.6e-133;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

RESULT 4
AAB36871
ID AAB36871 standard; protein; 348 AA.
XX
AC AAB36871;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation protein.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-00834497.
XX
PR 04-APR-1996; 96US-00630912.
PR 16-APR-1996; 96US-00632673.
PR 23-MAY-1996; 96US-00652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR N-PSDB; AAC68427.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful for
PT treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload.
XX
PS Claim 3; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene
XX
SQ Sequence 348 AA;

Query Match 100.0%; Score 1520; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.1e-133;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

RESULT 5
AAW94295
ID AAW94295 standard; peptide; 276 AA.
XX
AC AAW94295;
XX
DT 27-APR-1999 (first entry)
XX
DE Wild-type HFE polypeptide sequence.
XX
KW HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW transfusion; protein replacement therapy; hereditary hemochromatosis;
KW transferrin receptor; iron deficiency; anemia.

XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "indicated in the sequence listing as Arg"
XX
PN WO9856814-A1.
XX
PD 17-DEC-1998.
XX
PF 12-JUN-1998; 98WO-US012436.
XX
PR 13-JUN-1997; 97US-00876010.
XX
PA (PROG-) PROGENITOR INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 1999-080886/07.
XX
PT New treatment of an iron overload disease - comprises use of HFE
PT polypeptides provided in a complex with full length, wild type human
PT (2m), useful in protein replacement therapy.
XX
PS Claim 1; Page 13; 36pp; English.
XX
CC The present sequence represents a wild-type HFE polypeptide. The HFE
CC polypeptides (AAW94295-297) provided in a complex with full length, wild
CC type human beta-2-microglobulin (beta2m) form compositions in the
CC treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC other iron overload conditions resulting from secondary causes (e.g.
CC repeated transfusions). Data regarding the structure and function
CC correlations of HFE polypeptides is useful in designing drugs that
CC modulate the HFE gene and HFE activity. The polypeptides are also useful
CC in protein replacement therapy for individuals possessing a defective HFE
CC gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the polypeptides
CC are also useful in treating primary and secondary iron overload diseases.
CC The modulators of the transferrin receptor are useful in treating iron
CC deficiency conditions such as anemia, and in modulating the amount of
CC iron transported into a cell. The HFE polypeptides provide a molecular
CC basis for the relationship between HFE and iron metabolism, which enables
CC treatment of iron overload and deficiency diseases
XX
SQ Sequence 276 AA;

Query Match 99.5%; Score 1513; DB 2; Length 276;
Best Local Similarity 99.6%; Pred. No. 7.1e-133;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

RESULT 6

ABG72685
ID ABG72685 standard; protein; 276 AA.
XX
AC ABG72685;
XX
DT 05-MAR-2003 (first entry)
XX
DE Human haemochromatosis (HFE) mature protein.
XX
KW Human; haemochromatosis; HFE; hereditary haemochromatosis;
KW iron overload disease; iron deficiency disease; Beta2-microglobulin;
KW Beta2m; transferrin receptor; anaemia.
XX
OS Homo sapiens.
XX
PN US6391852-B1.
XX
PD 21-MAY-2002.
XX
PF 12-JUN-1998; 98US-00094964.
XX
PR 13-JUN-1997; 97US-00876010.
XX
PA (BIRA) BIO-RAD LAB INC.
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR WPI; 2003-155377/15.
XX
PT Method of treating an iron overload disease comprises administration of a
PT soluble complex comprising a 276 amino acid HFE polypeptide and a full
PT length, wild-type human beta2m.
XX
PS Claim 1; Col 1; 17pp; English.
XX
CC The invention relates to a method of treating an iron overload disease
CC comprising administration of a soluble complex comprising a 276 amino
CC acid mature HFE (hereditary haemochromatosis gene protein) polypeptide
CC (ABG72685-ABG72687) and a full length, wild-type human beta2m (beta2-
CC microglobulin). In a HeLa cell based assay, binding and uptake of ^1^2^5I
CC -transferrin in the presence of purified H63D-HFE/beta2m heterodimers was
CC determined. At a concentration of 250 nM H63D-HFE/ beta2m heterodimers,
CC the transferrin receptor (TfR) displayed a KD for transferrin of 28 nM.
CC At the same concentration of normal HFE/beta_2m heterodimers, TfR
CC displayed a KD for transferrin of 40 nM. In the absence of any
CC HFE/beta_2m heterodimers, TfR displayed a KD for transferrin of 7nM. It
CC was observed that H63D-HFE/beta_2m heterodimers were 30-40 % less
CC efficient in decreasing TfR affinity for transferrin compared to wild-
CC type HFE. The method is useful for treating iron overload diseases and
CC iron deficiency e.g. anaemia. The present sequence is wild-type mature
CC HFE
XX
SQ Sequence 276 AA;

Query Match 99.5%; Score 1513; DB 6; Length 276;
Best Local Similarity 99.6%; Pred. No. 7.1e-133;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 7
ABU62092
ID   ABU62092 standard; protein; 276 AA.
XX
AC   ABU62092;
XX
DT   01-OCT-2003  (first entry)
XX
DE   HFE mutant polypeptide #1 useful for treating iron diseases.
XX
KW   Iron overload disease; iron deficiency disease; HFE polypeptide;
KW   beta2 microglobulin; beta2m; hereditary haemochromatosis; HH; anaemia;
KW   protein replacement therapy; defective HFE gene; human; antianaemic;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2003073627-A1.
XX
PD   17-APR-2003.
XX
PF   04-MAR-2002; 2002US-00092404.
XX
PR   13-JUN-1997;   97US-00876010.
PR   12-JUN-1998;   98US-00094964.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Feder JN,  Bjorkman PJ,  Schatzman RC;
XX
DR   WPI; 2003-567313/53.
XX
PT   Treating an iron overload disease (e.g. hemochromatosis) or an iron
PT   deficiency disease (e.g. anemia), comprises administering to a patient an
PT   HFE polypeptide and full-length, wild type human beta-2 microglobulin.
XX
PS   Claim 3; Page 1; 14pp; English.
XX
CC   The present invention relates to a method for treating iron overload
CC   diseases and iron deficiency diseases. The method comprises administering
CC   to a patient an HFE polypeptide. The HFE polypeptide is provided in a
CC   complex with full-length, wild type human beta2 microglobulin (beta2m).
CC   The method and HFE polypeptide are useful for diagnosing or treating an
CC   iron overload disease (e.g. hereditary haemochromatosis, HH) or an iron
CC   deficiency disease (e.g. anaemia). The HFE polypeptide is also useful in
CC   protein replacement therapy for individuals having a defective HFE gene.
CC   The present sequence represents an HFE polypeptide useful for treating
CC   iron diseases
XX
SQ   Sequence 276 AA;

  Query Match             99.5%;  Score 1513;  DB 6;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 7.1e-133;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCPALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 8
AAW36499
ID   AAW36499 standard; protein; 348 AA.
XX
AC   AAW36499;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene product.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis; gene therapy;
KW   prenatal screening; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 63
FT                   /note= "substituted by Asp in 24s2 mutant"
FT   Misc-difference 65
FT                   /note= "substituted by Cys in 24d7 variant"
FT   Misc-difference 282
FT                   /note= "substituted by Tyr in 24d1 mutant"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US006254.
XX
PR   04-APR-1996;   96US-00630912.
PR   16-APR-1996;   96US-00632673.
PR   23-MAY-1996;   96US-00652265.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Thomas WJ,  Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Tsuchihashi Z;
PI   Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   N-PSDB; AAT96690, AAT96691.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis and
PT   treatment of hereditary haemochromatosis disease.
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This polypeptide is the expression product of a novel human gene (see
CC   AAT96690) whose mutated form is associated with hereditary
CC   haemochromatosis (HH). A single mutation (24d1) in the HH gene appears
CC   responsible for the majority of HH disease. This comprises a G to A
CC   substitution that is present in 86% of affected chromosomes and in 4% of
CC   unaffected chromosomes. It results in a Cys to Tyr substitution in the
CC   encoded protein at a critical disulphide bridge important for secondary
CC   structure. The following are claimed: the 10825 bp genomic DNA sequence
CC   (I), a 1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a peptide
CC   product chosen from the HH gene product, its variants (24d1, 24d2 and
CC   24d7), or a peptide of at least 56 amino acid residues of these; an
CC   antibody produced using the peptide as an immunogen; a method to
CC   determine the presence or absence of the common HH gene mutation; an
CC   animal model for the HH disease; metal chelation agents, T-cell
CC   differentiation factors and therapeutic agents for the mitigation of
```

```
CC   injury due to oxidative process in vivo or mitigation of iron overload; a
CC   method for screening potential therapeutic agents for activity in
CC   connection with HH disease; an antisense oligonucleotide directed against
CC   a transcriptional product of a nucleic acid sequence as above; and
CC   oligonucleotides or pairs of oligonucleotides covering a range of
CC   nucleotides from (I), (Ia) or their variants, useful for detecting a
CC   polymorphism in the HH gene. The invention also relates to methods for
CC   screening for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy, protein-
CC   and antibody-based therapeutics, and small molecule therapeutics
XX
SQ   Sequence 348 AA;

  Query Match             99.5%;  Score 1513;  DB 2;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 9.6e-133;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 9
AAB19149
ID   AAB19149 standard; protein; 348 AA.
XX
AC   AAB19149;
XX
DT   19-FEB-2001  (first entry)
XX
DE   A human histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .22
FT                   /note= "signal peptide"
FT   Misc-difference 63
FT                   /note= "when nucleotide 187 is mutated to G, then this
FT                   residue is Asp"
FT   Misc-difference 65
FT                   /note= "when nucleotide 193 is mutated to T, then this
FT                   residue is Cys"
FT   Domain          80. .108
FT                   /note= "alphal domain"
FT   Misc-difference 93
FT                   /note= "when nucleotide 277 is mutated to C, then this
FT                   residue is Arg"
FT   Misc-difference 105
FT                   /note= "when nucleotide 314 is mutated to C, then this
FT                   residue is Thr"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US007982.
XX
PR   26-MAR-1999;   99US-00277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   N-PSDB; AAA96769.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic acid.
XX
PS   Disclosure; Page 3; 55pp; English.
XX
CC   The present sequence represents a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC) non-
CC   classical class I gene located on chromosome 6p. Mutations in the gene
CC   lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of a
CC   mutation in exon 2 or an intron of a HFE gene or protein. The mutation is
CC   not a C to G missense mutation at nucleotide 187 of the sequence given in
CC   A96769 (Genbank Accession number U60319). The presence of the mutation
CC   indicates the disorder or the genetic susceptibility to the disorder. The
CC   method is used to diagnose an iron disorder e.g. haemochromatosis, or a
CC   genetic susceptibility to develop it
XX
SQ   Sequence 348 AA;

  Query Match             99.5%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 9.6e-133;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 10
AAB36869
ID   AAB36869 standard; protein; 348 AA.
XX
AC   AAB36869;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
```

OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-00834497.
XX
PR 04-APR-1996; 96US-00630912.
PR 16-APR-1996; 96US-00632673.
PR 23-MAY-1996; 96US-00652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR N-PSDB; AAC68425.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful for
PT treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload.
XX
PS Claim 1; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene
XX
SQ Sequence 348 AA;

Query Match 99.5%; Score 1513; DB 4; Length 348;
Best Local Similarity 99.6%; Pred. No. 9.6e-133;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

RESULT 11
AAB36872
ID AAB36872 standard; protein; 348 AA.
XX
AC AAB36872;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation protein.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload.
XX

OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-00834497.
XX
PR 04-APR-1996; 96US-00630912.
PR 16-APR-1996; 96US-00632673.
PR 23-MAY-1996; 96US-00652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR N-PSDB; AAC68428.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful for
PT treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload.
XX
PS Claim 4; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene
XX
SQ Sequence 348 AA;

Query Match 99.3%; Score 1509; DB 4; Length 348;
Best Local Similarity 99.6%; Pred. No. 2.3e-132;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

RESULT 12
AAU80035
ID AAU80035 standard; protein; 438 AA.
XX
AC AAU80035;
XX
DT 15-JUL-2002 (first entry)
XX
DE Beta 2 microglobulin (beta2M)/HFE monochain.
XX
KW Human; beta 2 microglobulin; beta2M/HFE monochain; HFE; ischaemia;
KW iron absorption regulator; intracellular iron absorption; lung injury;
KW haemochromatosis; transfusion; thalassaemia; haemolytic anaemia;

```
KW   chronic infection; transferrin receptor; TfR; brain tumour; cancer;
KW   oxidative stress disorder; tissue damage; vascular disease; inflammation;
KW   atherosclerosis; autoimmune disease; inflammatory condition.
XX
OS   Homo sapiens.
XX
PN   WO200224929-A2.
XX
PD   28-MAR-2002.
XX
PF   24-SEP-2001; 2001WO-US029873.
XX
PR   22-SEP-2000; 2000US-0234843P.
XX
PA   (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA   (MCIN/) MCINNIS P.
XX
PI   Ehrlich R, Rotem-Yehudar R, Laham N;
XX
DR   WPI; 2002-383192/41.
DR   N-PSDB; ABK49917.
XX
PT   Soluble beta 2 microglobulin/HFE monochain useful for treating iron-
PT   overload conditions e.g. thalassemia and chronic infections, comprises
PT   human beta 2 microglobulin linked to alpha domains of HFE by a linker
PT   peptide.
XX
PS   Example 2; Fig 2; 77pp; English.
XX
CC   The invention relates to a soluble polypeptide (I) of beta 2
CC   microglobulin (beta2m)/HFE monochain comprising human beta2m (or its
CC   analogue or active fragment), linked to alpha1-alpha3 domains of human
CC   HFE (a central regulator of iron absorption; undefined), or its analogue
CC   or active fragment, by a flexible linker peptide, or a functional
CC   derivative or salt of (I). (I) is useful for reducing intracellular iron
CC   absorption in patients having hereditary haemochromatosis, transfusions,
CC   thalassaemias, haemolytic anaemia or chronic infections, and for
CC   delivering a therapeutic to cells that over-express transferrin receptor
CC   (TfR) which are preferably lymphocytes or leukocytes, across the blood-
CC   brain barrier. (I) is further useful for treating brain tumour. (I) is
CC   also useful for treating oxidative stress disorders resulting in tissue
CC   damage e.g. vascular diseases, inflammation, atherosclerosis, lung
CC   injury, ischaemia, etc. A DNA molecule (II) encoding (I) is useful as a
CC   platform for drug delivery of therapeutic use for cancer, autoimmune
CC   diseases and inflammatory conditions. The monochain manifests specific
CC   characteristics advantageous for drug delivery systems. It is a soluble,
CC   stable and fully conformed protein. It binds specifically to transferrin
CC   receptor (TfR) and therefore targets cells that over-express this
CC   receptor. It is continuously internalised by the target cells, thus
CC   enabling efficient drug delivery. It dissociates from the receptor in the
CC   cells, minimising side effects. It negatively regulates iron absorption,
CC   reducing growth of undesired cells and preventing lymphocyte activation.
CC   It is not diluted in the blood as is transferrin. It should not induce an
CC   immune response since it is a self non-polymeric protein and delivery of
CC   drugs via monochain is expected to overcome drug-resistance since it is a
CC   natural TfR-binding protein. The present sequence represents the amino
CC   acid sequence of beta2m/HFE monochain
XX
SQ   Sequence 438 AA;

  Query Match             99.2%;  Score 1508;  DB 5;  Length 438;
  Best Local Similarity   99.6%;  Pred. No. 3.8e-132;
  Matches  274;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        135 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 194

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 254

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 314

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 374

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
              ||||||||||||||||||||||||||||||||||
Db        375 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 409


RESULT 13
ABU62091
ID   ABU62091 standard; protein; 276 AA.
XX
AC   ABU62091;
XX
DT   01-OCT-2003  (first entry)
XX
DE   HFE polypeptide useful for treating iron diseases.
XX
KW   Iron overload disease; iron deficiency disease; HFE polypeptide;
KW   beta2 microglobulin; beta2m; hereditary haemochromatosis; HH; anaemia;
KW   protein replacement therapy; defective HFE gene; human; antianaemic;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2003073627-A1.
XX
PD   17-APR-2003.
XX
PF   04-MAR-2002; 2002US-00092404.
XX
PR   13-JUN-1997;   97US-00876010.
PR   12-JUN-1998;   98US-00094964.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Feder JN, Bjorkman PJ, Schatzman RC;
XX
DR   WPI; 2003-567313/53.
XX
PT   Treating an iron overload disease (e.g. hemochromatosis) or an iron
PT   deficiency disease (e.g. anemia), comprises administering to a patient an
PT   HFE polypeptide and full-length, wild type human beta-2 microglobulin.
XX
PS   Claim 1; Page 1; 14pp; English.
XX
CC   The present invention relates to a method for treating iron overload
CC   diseases and iron deficiency diseases. The method comprises administering
CC   to a patient an HFE polypeptide. The HFE polypeptide is provided in a
CC   complex with full-length, wild type human beta2 microglobulin (beta2m).
CC   The method and HFE polypeptide are useful for diagnosing or treating an
CC   iron overload disease (e.g. hereditary haemochromatosis, HH) or an iron
CC   deficiency disease (e.g. anaemia). The HFE polypeptide is also useful in
CC   protein replacement therapy for individuals having a defective HFE gene.
CC   The present sequence represents an HFE polypeptide useful for treating
CC   iron diseases
XX
SQ   Sequence 276 AA;

  Query Match             99.1%;  Score 1506;  DB 6;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 3.2e-132;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          51 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy         121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy         181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
                  ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         181 DQQVPPLVKVTHHVTSSVTTLRCPALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy         241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                  ||||||||||||||||||||||||||||||||||
Db         241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 14
AAB36870
ID   AAB36870 standard; protein; 348 AA.
XX
AC   AAB36870;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-00834497.
XX
PR   04-APR-1996;   96US-00630912.
PR   16-APR-1996;   96US-00632673.
PR   23-MAY-1996;   96US-00652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68426.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful for
PT   treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload.
XX
PS   Claim 2; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene
XX
SQ   Sequence 348 AA;

  Query Match             98.8%;  Score 1502;  DB 4;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1e-131;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy           1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
                 ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy          61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy         121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy         181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy         241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
                  |||||||||||||||||| ||||||||||||||||
Db         263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 15
AAW94297
ID   AAW94297 standard; peptide; 276 AA.
XX
AC   AAW94297;
XX
DT   27-APR-1999  (first entry)
XX
DE   HFE mutant (H111A/H145A-HFE) polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia; mutant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
FT   Misc-difference 89
FT                   /label= H111A
FT                   /note= "wild type His (of the mature protein sequence) is
FT                   replaced by Ala"
FT   Misc-difference 123
FT                   /label= H145A
FT                   /note= "wild type His (of the mature protein sequence) is
FT                   replaced by Ala"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US012436.
XX
PR   13-JUN-1997;   97US-00876010.
XX
PA   (PROG-) PROGENITOR INC.
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI   Feder JN,  Bjorkman PJ,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy.
XX
PS   Claim 5; Page 15; 36pp; English.
XX
CC   The present sequence represents a H111A/H145A-HFE mutant polypeptide. The
CC   HFE polypeptides (AAW94295-297) provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
```

CC in protein replacement therapy for individuals possessing a defective HFE
CC gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the polypeptides
CC are also useful in treating primary and secondary iron overload diseases.
CC The modulators of the transferrin receptor are useful in treating iron
CC deficiency conditions such as anemia, and in modulating the amount of
CC iron transported into a cell. The HFE polypeptides provide a molecular
CC basis for the relationship between HFE and iron metabolism, which enables
CC treatment of iron overload and deficiency diseases
XX
SQ Sequence 276 AA;

Query Match 98.2%; Score 1493; DB 2; Length 276;
Best Local Similarity 98.9%; Pred. No. 5.2e-131;
Matches 273; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              |||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

Search completed: May 4, 2004, 11:35:02
Job time : 49.3333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 4, 2004, 11:36:43 ; Search time 36 Seconds
(without alignments)
2125.120 Million cell updates/sec

Title: US-10-092-404-2
Perfect score: 1520
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1520	100.0	276	14	US-10-092-404-2	Sequence 2, Appli
2	1520	100.0	348	14	US-10-138-888-6	Sequence 6, Appli
3	1513	99.5	276	14	US-10-092-404-1	Sequence 1, Appli
4	1513	99.5	348	10	US-09-981-606-2	Sequence 2, Appli
5	1513	99.5	348	14	US-10-138-888-2	Sequence 2, Appli
6	1509	99.3	348	14	US-10-138-888-8	Sequence 8, Appli
7	1508	99.2	348	14	US-10-138-888-78	Sequence 78, Appl
8	1502	98.8	348	14	US-10-138-888-4	Sequence 4, Appli
9	1493	98.2	276	14	US-10-092-404-3	Sequence 3, Appli
10	542.5	35.7	341	15	US-10-143-822-1	Sequence 1, Appli
11	523	34.4	361	14	US-10-138-888-22	Sequence 22, Appl
12	514	33.8	92	13	US-10-016-634A-120	Sequence 120, App
13	514	33.8	365	14	US-10-138-888-23	Sequence 23, Appl
14	506	33.3	280	14	US-10-073-300-6	Sequence 6, Appli
15	506	33.3	415	14	US-10-073-300-5	Sequence 5, Appli
16	506	33.3	510	12	US-10-108-511-5	Sequence 5, Appli
17	492	32.4	298	14	US-10-205-823-40	Sequence 40, Appl
18	492	32.4	298	14	US-10-205-823-42	Sequence 42, Appl
19	492	32.4	298	14	US-10-177-293-23	Sequence 23, Appl
20	492	32.4	326	12	US-10-380-880-7	Sequence 7, Appli
21	490.5	32.3	379	12	US-10-210-172-160	Sequence 160, App
22	490.5	32.3	379	15	US-10-093-463-78	Sequence 78, Appl
23	487.5	32.1	389	12	US-10-108-511-2	Sequence 2, Appli
24	477	31.4	542	14	US-10-015-535-32	Sequence 32, Appl
25	477	31.4	542	14	US-10-015-535-34	Sequence 34, Appl
26	475	31.2	542	14	US-10-015-535-36	Sequence 36, Appl
27	474	31.2	540	14	US-10-015-535-22	Sequence 22, Appl
28	474	31.2	541	14	US-10-015-535-28	Sequence 28, Appl
29	474	31.2	542	14	US-10-015-535-24	Sequence 24, Appl
30	474	31.2	542	14	US-10-015-535-26	Sequence 26, Appl
31	465	30.6	364	15	US-10-093-463-80	Sequence 80, Appl
32	462	30.4	362	12	US-10-257-021-82	Sequence 82, Appl
33	459	30.2	371	12	US-10-210-172-156	Sequence 156, App
34	459	30.2	371	15	US-10-085-198-72	Sequence 72, Appl
35	458	30.1	362	12	US-09-819-371-4	Sequence 4, Appli
36	457	30.1	274	12	US-09-819-371-5	Sequence 5, Appli
37	448	29.5	332	9	US-09-870-521-3	Sequence 3, Appli
38	445	29.3	540	14	US-10-015-535-30	Sequence 30, Appl
39	444	29.2	334	9	US-09-870-521-4	Sequence 4, Appli
40	431.5	28.4	389	15	US-10-085-198-70	Sequence 70, Appl
41	415.5	27.3	452	12	US-10-210-172-152	Sequence 152, App
42	415.5	27.3	452	15	US-10-085-198-68	Sequence 68, Appl
43	399	26.2	421	12	US-10-210-172-174	Sequence 174, App
44	399	26.2	421	15	US-10-138-588-32	Sequence 32, Appl
45	379	24.9	284	15	US-10-104-047-3648	Sequence 3648, Ap

ALIGNMENTS

RESULT 1
US-10-092-404-2
; Sequence 2, Application US/10092404
; Publication No. US20030073627A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; Bjorkman, Pamela J.
; Schatzman, Randall C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
; AND IRON DEFICIENCY DISEASES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/092,404
; FILING DATE: 04-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,964
; FILING DATE: June 12, 1998
; APPLICATION NUMBER: 08/876,010
; FILING DATE: June 13, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0074-999
; TELECOMMUNICATION INFORMATION:

```
;             TELEPHONE: 650-493-4935
;             TELEFAX: 650-493-5556
;             TELEX: 66141 PENNIE
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 276 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-092-404-2

  Query Match            100.0%;  Score 1520;  DB 14;  Length 276;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-146;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-10-138-888-6
; Sequence 6, Application US/10138888
; Publication No. US20030148972A1
;   GENERAL INFORMATION:
;        APPLICANT: Thomas, Winston J.
;                   Drayna, Dennis T.
;                   Feder, John N.
;                   Gnirke, Andreas
;                   Ruddy, David
;                   Tsuchihashi, Zenta
;                   Wolff, Roger K.
;        TITLE OF INVENTION: Hereditary Hemochromatosis Gene
;        NUMBER OF SEQUENCES: 79
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10036-2711
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/138,888
;             FILING DATE: 02-May-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/834,497
;             FILING DATE: 04-APR-1997
;             APPLICATION NUMBER: US 08/652,265
;             FILING DATE: 23-MAY-1996
;             APPLICATION NUMBER: US 08/632,673
;             FILING DATE: 16-APR-1996
;             APPLICATION NUMBER: US 08/630,912
;             FILING DATE: 04-APR-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brian M. Poissant
;             REGISTRATION NUMBER: 28,462
;             REFERENCE/DOCKET NUMBER: 8907-095-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 790-9090
;             TELEFAX: (212) 869-8864
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 348 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-138-888-6

  Query Match            100.0%;  Score 1520;  DB 14;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-145;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy         61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy        121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy        181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy        241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
              ||||||||||||||||||||||||||||||||||||
Db        263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
US-10-092-404-1
; Sequence 1, Application US/10092404
; Publication No. US20030073627A1
;   GENERAL INFORMATION:
;        APPLICANT: Feder, John N.
;                   Bjorkman, Pamela J.
;                   Schatzman, Randall C.
;        TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
;                            DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;                            AND IRON DEFICIENCY DISEASES
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Pennie & Edmonds, LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New York
;             STATE: NY
;             COUNTRY: USA
;             ZIP: 10036-2811
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/092,404
;             FILING DATE: 04-Mar-2002
;             CLASSIFICATION: <Unknown>
```